

Best Local Similarity 64.9%; Pred. No. 5.8e-82;
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

```

QY 27 GPDTEAEDVDVPMWYQSSAKGIDYDKLIVRGSSKIDKELNRIERATGQRPHHFLR 86
Db 23 GGVQVEDEEDVETWETTKATGIDYDKLIVKGCRLDEILARERVRGHKASPLR 82
QY 87 RGIFSHRDMQVLDAYENKRPFLYTGRCSSAMHGHILPIFTKWLQDVNVLPI 146
Db 83 RGMFAHRDLAILDRKEGKPFYLYTGRCASSSLHGHILVPIFTKWLQDVNVLPI 142
QY 147 QMTDEKYLKMDLTLDAYDAVENAKDIACGFIDINKTIFISDLYMGSSGFYKNV 206
Db 143 QMTDEKFLMKMAYDEAKKMARENMKDIIISVGFDPKTFIFNNFDY--MCPPYENIVK 200
QY 207 IQKAVTNQVKGIFGFTDSDICGISFPALQAPSFNSFQIIRDRDIOCLIPCAIDQ 266
Db 201 IWKVNTNQAVILFEPEDCLGRAAPVAVAPCAAFSSFPQIFGRNDICLIPCAIDQ 260
QY 267 DPFYRMRDVAIRIGYKPKALHSTFPALQAGTQKMSASPNSIFLDTAKOIKTKV 325
Db 261 DPFYRMRDVAIRIGYKPKALHSTFPALQAGTQKMSASPNSIFLDTAKOIKTKV 319

```

RESULT 10

```

QY 0976M1 PRELIMINARY; PRT; 406 AA.
AC 0976M1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.
GN S70169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
KM EMBL: AP000981; BAB65126.1;
KW Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA; 47979 MW; 276CCD6506B94B3 CRC64;

```

Query Match 40.5%; Score 910.5; DB 17; Length 406;
Best Local Similarity 50.5%; Pred. No. 1.4e-71;
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

```

QY 35 EDFVDPMTVQTSKAG-IDYDKLIVRGSSKIDKELNRIERATGQRPHHFLRGI 92
Db 28 QDFVNTVPEV-----KGAVDYDKLIVGFTOKITSELKEKIKSIINDELHVMRLRDVFS 82
QY 93 HRDMNOVLDAVENKRPFLYTGRCSSAMHGHILPIFTKWLQDVNVLPIQMTDE 152
Db 83 HRDLVLVAKDQDGKGFYLYTGRCAPSL-GMHIGHILPIFTKWLQDKRNVALLYEITDE 141
QY 153 KYLMK-DLTLQAYGDAVENAKDIACGFIQDKNKTIFISDLYMGSSGFYKNVYK 211
Db 142 KFMNPEVTLDTQOMAYDNLDIIAVGFNDPKTFIFQDTEYI---RMVYIAIKIKAL 198
QY 212 TFNVKVGIFGFTDSDICGISFPALQAPSFNSFQIIRDRDIOCLIPCAIDDPYR 271
Db 199 TFEVRAITFGIDTSNIGITWYPALQIAPT-----MFEK---KCLIPAGIDDPYR 248

```

```

QY 272 MTRDVAIRIGYKPKALHSTFPALQAGTQKMSASPNSIFLDTAKOIKTKVKN 331
Db 249 LQRIAESLGYKKAQIHSKRLPLTGEGKSSQPTAIYLLDDPRTVARKIMTKAFS 308
QY 332 GGRPTIEHROFGNCIDVFSMYLTFPLE-DDDKLEIRKDYSGAMLTGELKALIEV 390
Db 309 GGPPTIELHRRYGGNPIDIVSFQWLYMFEFEDDNKIKIEEDYRSAGALLTELKQIL 368
QY 391 LQPLIAEQARKE 404
Db 369 LNDPLEHQRKEE 382

```

RESULT 11

```

QY 0972X0 PRELIMINARY; PRT; 386 AA.
AC 0972X0;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE TRYPTOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).
GN TRPS.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.R.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Noc H.P., Redder P., Schenk M.E., Theinault C., Tolstrup N.,
RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus p2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006677; AAK40778.1;
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PRO1039; TRNASYNTHTRP.
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ SEQUENCE 386 AA; 45448 MW; 3F4DA8CB80D219DF CRC64;

```

Query Match 40.4%; Score 907; DB 17; Length 386;
Best Local Similarity 48.6%; Pred. No. 2.6e-71;
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

```

QY 31 TEAEDVDVPMWYQSSAKG-IDYDKLIVRGSSKIDKELNRIERATGQRPHHFLRGI 89
Db 6 TMDPEFTVPEV-----KGAVDYDKLIVGFTOKITSELKQRIKLNLAGDL-HVMRLRVN 59
QY 90 FFHRDMNOVLDAVENKRPFLYTGRCSSAMHGHILPIFTKWLQDVNVLPIQMT 149
Db 60 FFHRDLVLVNDYEKSGFELYTGRCAPSL-GMHIGHILPIFTKWLQDKRNVALLYEIT 118
QY 150 DDEKYLK-DLTLQAYGDAVENAKDIACGFIQDKNKTIFISDLYMGSSGFYKNV 208
Db 119 DDEKYMNPETLDTQTRMAYDNLDIIAVGFNDPKTFIFQDTEYI---RMVYIAIK 175
QY 209 KHAVENOVKGFEGFTDSDICGISFPALQAPSFNSFQIIRDRDIOCLIPCAIDDP 268
Db 176 KILTFEVRATFGIDASSNIGILIFPALQIAPT-----MFEK---KCLIPAGIDDP 225
QY 269 YFRMRDVAIRIGYKPKALHSTFPALQAGTQKMSASPNSIFLDTAKOIKTKV 328
Db 226 YMRLORDIAESLGYKKAQIHSKRLPLTGEGKSSQPTAIYLLDDPRTVARKIMTK 285
QY 329 AFSGRPTIEHROFGNCIDVFSMYLTFLEDD-DKLEIRKDYSGAMLTGELKAL 387
Db 286 AFSGGPTIELHRRYGGNPIDIVSFQWLYMFEFEDDNKIKIEEDYRSAGALLTELKQIL 345

```

RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003682; AAG22136.1; -
 DR FlyBase: FBgn0010803; Aats-trp.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1b.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_I; 1.
 SQ SEQUENCE 430 AA; 47971 MW; 2338EECC69E979F CRC64;

Query Match 61.4%; Score 1378; DB 5; Length 430;
 Best Local Similarity 61.4%; Pred. No. 1.3e-112;
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

QY 10 DYKADCEPGNAPTSNHCPRD-----TEAE-----EDFDPPTVOTSSAKGIDYD 54
 DB 3 DTKEVVEGVEALTLNKPDAPEVETGDAQOAGATAPTDVDDVPMNVASSNAGVDYD 62
 QY 55 KLIVFGSSSKIDKELINRIERATGPRPHFRLRGIFPSHRMNOVLDAEYKKKPFYLYTG 114
 DB 63 KLIRFGSSSKIDELIAFEKITGRPAHFRGMPFSHRDLITLIREGKPFYLYTG 122
 QY 115 RGPSEAMHVGHLIPFTFKWLDVFNVLVYQMTDEKYLMDLTDQAYGDAVENAKD 174
 DB 123 RGPSSGLHVGHVLPITMTKWLQETFDVPLVYQMTDEKYLMDLTDQAYGDAVENAKD 182
 QY 175 ILAGGPDINKFTISDLDYMGSSGFYKNVYKIOKHTFNQYKIFGGTSDCIGKISFP 234
 DB 183 IVAIGFVNKFTIENNEFEVKGCPAMYONIRIOKCVTFNOKGIFGGSDDIIGKIGFP 242
 QY 235 AIOGAPFSNFPDIFRDRDIOCLIPCAIDDDPYFRMTRAVARIGYKPKALLHSFPF 294
 DB 243 AAOGAPALISFPPIFGNR-KVHCLIPCAIDDDPYFRMTRAVARIGYKPKALLHSFPF 301
 QY 295 ALGAGQTKMSASDPNSSIFLTDPAKOIKTYKNKHAFGSGRDTIEHROFGNCVDVSFM 354
 DB 302 ALGAGTKMSASDPNSSAVYLTDPKQIKNKINKYAFSGRVSVEHNRKGLGPEVDVSYQ 361
 QY 355 YLTFLEDDDKLEQIRKDYTGAMLTGELKALLLEVLOPLAEQARKKETDELVEKFM 414
 DB 362 LKFELEDDDALEEVAVSNGEMLTGEIKLAVETLPIVEHQAAKKLITDELVKYF 421
 QY 415 TPKLSF 421
 DB 422 ELRLPKF 428

RESULT 8
 Q9SR15 PRELIMINARY; PRT; 402 AA.
 AC Q9SR15;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
 DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.
 GN 7018.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Beutlo M., Creasy T.H., Haas B.,
 RA Rensing C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Niernan W.C., Fraser C.M.;
 RA "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011437; AAF04890.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.

DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_I; 1.
 DR Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75EE506CD15 CRC64;

Query Match 57.9%; Score 1301; DB 10; Length 402;
 Best Local Similarity 60.3%; Pred. No. 6.8e-106;
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

QY 29 DATEAE--EDFVDVPTVOTSSAKGIDYDKLIVFGSSSKIDKELINRIERATGPRPHFLR 86
 DB 7 DERAESESEQVNVNWEVSARKGKIDYDKLIDKRGCCQLDSESLDRVQRLSRQPHFLR 66
 QY 87 RGIFFSHRDMNOVLDAEYKKKPFYLYTGSGSSSEAMHVGHLIPFTFKWLDVFNVLVY 146
 DB 67 RSVFAHRDFENELDAEYRGDKFYLYTGRGSSSALHGLIPMTFTYLOEAKFVPLVY 126
 QY 147 QMTDEKYLMDLTDQAYGDAVENAKDIACGFDINKFTISDLDYMGSSGFYKNVYK 206
 DB 127 QLTDEKSIKMLNLSVEESQRLARENARDIACGFDVTKFTIFSPPDIVG--GAFYKNVYK 184
 QY 207 IQKHTFNQYKIFGFTSDCIGKISFPALQAPFSNFPDIFRDRDIOCLIPCAIDQ 266
 DB 185 VGKCVTLNKAKGIFGSESDPIALSPFPVQAVSPSSFPHLRPGKONLCLIPCAIDQ 244
 QY 267 DPTFRMTRDVAPRIGYKPKALLHSFPALOGAOTKMSASDPNSSIFLTDPAKOIKTYN 326
 DB 245 DPTFRMTRDVAPRIGYKSPALISTEFPALOGENGKMSASDPNSAIYVTDKAKIKKNIN 304
 QY 327 KHAFGSGRDTIEHROFGNCVDVSFMVLFPEDDDKLEQIRKDYTGAMLTGELKKA 386
 DB 305 RYAFSGGDSITEKRELGANLEVDIPKYLSFLEDDSELHINKETGEGMLTGEYKRR 364
 QY 387 LIEVLOPLAEHQARKKEVTDEIVKFTPRKLSFDPQ 424
 DB 365 LIEVLTIEVHRHRAAVALTDEMVAFAVAPRLPSMFE 402

RESULT 9
 Q9UIR2 PRELIMINARY; PRT; 324 AA.
 AC Q9UIR2;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)
 DE Y80D3A.1 PROTEIN.
 GN Y80D3A.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Harris B.R.;
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=99069613; Pubmed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: AL132853; CAB60439.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 SQ SEQUENCE 324 AA; 36289 MW; 6E687E6D420ECC12 CRC64;

Query Match 45.7%; Score 1027; DB 5; Length 324;

QY	115	RGSSSEAMHVGHLDPFTTKMLQDVENVPLVYQMTQDDEKYLKMDLTLDQAYGDAVENAKD	174
DB	123	RGPSSESLHVGHLDPFTTKMLQDDEKYLKMDLTLDQAYGDAVENAKD	182
QY	175	IIACGFDFNKFIFESDDLYDMGSSGFEKNNVYIKQHTYFNPQVKIIFGFTPSDICKITSP	234
DB	183	IYALISFDVNAKFFIFNNLEFEVKKCAMTONILIRKCVTFPNQVKIIFGFGSDITKIGFP	242
QY	235	AIQAAFSFNSPFOITFRRTDIOCLIPCADIDQDYFRTRTQVAPRIGYPRPALIHSTFP	294
DB	243	AAQAAPALISTFPEFIFGNR-KVCHCLIPCADIDQDYFRTRTQVAPRIGYPRPALIHSTFP	301
QY	295	ALQAGQATMSASDRSSLFILDTQAKQIKTKVNAKFAFGSDPTIEHHQFGNCDVDVSEF	354
DB	302	ALQAGKATMSASDRSSAYVLLDTPEKQIKTKVNAKFAFGSDPTIEHHQFGNCDVDVSEF	361
QY	355	YLTFFLEDDDLKLEQIRKQYKTSAGMLTGELKALLEVLPILAEIQARKKEVTPDVEFM	414
DB	362	LKFFLEDDDLKLEQIRKQYKTSAGMLTGELKALLEVLPILAEIQARKKEVTPDVEFM	421
QY	415	TPRKLSF 421	
DB	422	ELRPLKF 428	
RESULT 6			
Q904Y0			
AC	Q904Y0	PRELIMINARY; PRT; 420 AA.	
DT	01-MAY-2000 (TREMBLREL_13, Created)		
DT	01-MAY-2000 (TREMBLREL_13, Last sequence update)		
DT	01-JUN-2001 (TREMBLREL_17, Last annotation update)		
DE	TRIPTOPHANYL-TRNA SYNTHETASE (FRAGMENT).		
GN	AATS-TRP OR CG9735.		
OS	Drosophila melanogaster (Fruit fly).		
CC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
CC	Pterygota; Neoptera; Endopterygota; Diptera; Braachyera; Muscomorpha;		
CC	Ephydroidea; Drosophilidae; Drosophila.		
OK	NCBI_TaxID=7227;		
LN	[1]		
RP	SEQUENCE FROM N.A.		
RP	Medline=99250164; Pubmed=10233165;		
RA	Seshiah P., Andrew D.J.;		
RT	"MS-55D: A tryptophanyl-tRNA synthetase expressed to high levels in		
RT	the developing drosophila salivary gland."		
RL	Mol. Biol. Cell 10:1595-1608(1999).		
DR	EMBL; AF125157; AAF20167.1; .		
DR	FLYBase; FBgn0010803; Aats-trp.		
DR	InterPro: IPR002305; tRNA-synt_1b.		
DR	InterPro: IPR001412; tRNA-synt_1b.		
DR	InterPro: IPR002306; tRNA-synt_trp.		
DR	Pfam; PF00579; tRNA-synt_1b; 1.		
DR	PRINTS; PR01039; TRNASYNTHTRP.		
DR	PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.		
KW	Aminoacyl-tRNA synthetase.		
FT	NON TER		
SO	SEQUENCE 420 AA; 46975 MW; 4AF708BAF426ABE6 CRC64;		
Query Match			
Best Local Similarity 62.74; Pred. No. 9.1e-113;			
Matches 257; Conservative % 62; Mismatches 90; Indels 1; gaps			
QY	12	KADRGKAPRATNSINGRPTAEEDVDYPRWYQSSAKGIDYKLLYLRFGSSKIDKRLIN	71
DB	10	KDAEVRVETGTDAAQEGASATPTEDVDYDPMVNAASNAGVDYDLKIRFGSSKIDDELLA	69
QY	72	RIERATQGRPHHFLRQGFESHRDMQNLDAVEKKKPYLYTGRGSSSEAMHVGHLPTI	131
DB	70	RPEKTGRPHHFLRQGFESHRDMQNLTLRLRQGRKPYLYTGRGSSSEAMHVGHLPTI	129
QY	132	FTTKLQDVENVPLVYQMTQDDEKYLKMDLTLDQAYGDAVENAKDIIACGFDFNKFIFES	191

Db	130	MTKWLQELFFDVLVLIQLTDEKDTLMLKAKEDAIKLRERAKOIVAIIGDVNKTFFFNLU	189
Qy	192	DYMGSSSEFYANVYKICKHNTFNVOKGIGFETFDSDCIKISFPAIOAASFSNPFOIR	251
Db	190	EFVGCPCAMYNQNIIRICQCVTFENVKGIIFGCGSDIIGKGFPAQAQAAPASISTPFPIG	249
Qy	252	DRTDIOCLIPCAIDODDPFMRTRDVAAPRIGPKPKALIHSTFPALOGAOTKMSADPNSS	311
Db	250	NR-KVHCILIPCAIDODDPFMRTRDVAAPRIGPKPKALIHSTFPALOGAOTKMSADPNSSA	308
Qy	312	IFLDTAOKITKYNKAHFASSGSGRTIEHHPGNCNDVDSFMYLNFLEDDOKLIDK	371
Db	309	VYLDTPKQIKNKINKKIAFSGSGRTVEHHRKLGVPVWDVSYOLLFLPEDDAKLEVRV	368
Qy	372	DYTSAGMLTGLKALIEVLQPLTAEHQARKKETDIEVYEFMTPKRISF	421
Db	369	AVSKEMLTGELKALVETLPIPIEQHQAARKLITDEVLDKYELARPIKF	418

RESULT 7

Q9VHG2 PRELIMINARY: PRT: 430 AA.

Q9VHG2

AC 01-MAY-2000 (TRMBLrel, 13, Created)

DT 01-MAY-2000 (TRMBLrel, 13, Last sequence update)

DT 01-JUN-2001 (TRMBLrel, 17, Last annotation update)

DE AATS-TRP PROTEIN.

GN AATS-TRP OR CG9735.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota, Metazoa: Arthropoda, Tracheata: Hexapoda: Insecta:

OC Pterygota, Neoptera: Endopterygota, Diptera: Brachycera: Muscomorpha:

OC Ephydroidea: Drosophilidae: Drosophila.

OX NCBI_TaxID=7227;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos K., Delcher A., Deng Z., Duan A.D., Dewey C., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Mayano-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz G., Ferriera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glöckner A., Gong F., Gottell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Havsey D., Helman T.J., Hernandez J.R., Houck C.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigbam C.,

RA Jaitani M., Kalush F., Karen G.H., Ke Z., Kemison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshireli A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Plittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stempleton M., Strong R., Sun E.,

RA Svrtkash R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmberg L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawanji H., Kotsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK007754; BAB5235.1; -
 DR MGD: MGI:104630; Marts.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; tRNA-syntHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244CA2A CRC64;

Query Match 72.5%; Score 1629; DB 11; Length 329;
 Best Local Similarity 92.1%; Pred. No. 7.7e-135;
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 96 MNOVDAVENKKPFYLYTGRGPSSSEAMVHGLPIFTKMLQDVFNPLVLOMTDDEKYL 155
 DB 1 MNOIIDAENKKPFYLYTGRGPSSSEAMVHGLPIFTKMLQDVFNPLVLOMTDDEKYL 60
 QY 156 WKDLTLDOAYGDAVENAKDIIACGFDINKTEIFSDLDYMGSSGFRYKNVRIQKHVTFNQ 215
 DB 61 WKDLTLDOAYGDAVENAKDIIACGFDINKTEIFSDLDYMGSSGFRYKNVRIQKHVTFNQ 120
 QY 216 VKGIFGFDSOCIGKISFPAIQAPSFNSFPQIFRDRTDIOCLIPCAIDDDPYFRMTRD 275
 DB 121 VKGIFGFDSOCIGKISFPAIQAPSFNSFPQIFRDRTDIOCLIPCAIDDDPYFRMTRD 180
 QY 276 VAPRIGYKPKPLHSTFEPALOGAOTKMSADPNSSIFLDTAKOIKKVKHAFSGGRD 335
 DB 181 VAPRIGYKPKPLHSTFEPALOGAOTKMSADPNSSIFLDTAKOIKKVKHAFSGGRD 240
 QY 336 TIEEHROFGNCQDVSEFMVLTFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 395
 DB 241 TIEEHROFGNCQDVSEFMVLTFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 300
 QY 396 AEHQARRKEVTDVYKEMTPRKISFDFQ 424
 DB 301 AEHQARRKEVTDVYKEMTPRKISFDFQ 329

RESULT 4

QY 070184 PRELIMINARY; PRT; 305 AA.
 ID 070184
 AC 070184
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE TRYPHOPHAN-TRNA SYNTHETASE (FRAGMENT).
 OS Cavia Porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae; Cavia.
 OX NCBI_Taxid=10141;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY; TISSUE=SPLEEN;
 RA Yang D., Goto R., Watanabe N., Kobayashi Y.;
 RT "Identification and Cloning of Genes Whose Expressions are Elevated
 RT during DMCB-Induced Guinea Pig Skin Delayed-type Hypersensitivity
 RT Reaction."
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB012222; BAA25288.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; tRNA-syntHTRP.
 KW Aminoacyl-tRNA synthetase.
 FT NON_TER 1
 SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 68.4%; Score 1537; DB 11; Length 305;
 Best Local Similarity 94.4%; Pred. No. 8.4e-127;
 Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 120 EAMVGHGLPIFTKMLQDVFNPLVLOMTDDEKYLKDLTLDOAYGDAVENAKDIIACG 179
 DB 1 EAMVGHGLPIFTKMLQDVFNPLVLOMTDDEKYLKDLTLDOAYGDAVENAKDIIACG 60
 QY 180 FDIKKTFFSDLDYMGSSGFRYKNVRIQKHVTFNQKGIFFGFDSCIGKISFPAIQAA 239
 DB 61 FDIKKTFFSDLDYMGSSGFRYKNVRIQKHVTFNQKGIFFGFDSCIGKISFPAIQAA 120
 QY 240 PSFNSFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPKPLHSTFEPALOGA 299
 DB 121 PSFNSFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPKPLHSTFEPALOGA 180
 QY 300 QTKMSADPNSSIFLDTAKOIKKVKHAFSGGRDTIEEHROFGNCQDVSEFMVLTFF 359
 DB 181 QTKMSADPNSSIFLDTAKOIKKVKHAFSGGRDTIEEHROFGNCQDVSEFMVLTFF 240
 QY 360 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDVYKEMTPRK 419
 DB 241 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDVYKEMTPRK 300
 QY 420 SFDFQ 424
 DB 301 SFDFQ 305

RESULT 5

0904Y1

QY 0904Y1 PRELIMINARY; PRT; 430 AA.
 ID 0904Y1
 AC 0904Y1
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE TRYPHOPHAN-TRNA SYNTHETASE.
 GN AATS-TRP OR CG9735.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC MEDLINE=99250164; PubMed=10233165;
 RA Seshalah P., Andrew D.J.;
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
 RT the developing Drosophila salivary gland."
 RL Mol. Biol. Cell 10:1595-1608(1999).
 DR EMBL: AF125156; AAR20166.1; -
 DR FlyBase: FBgn0010803; Aats-trp.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; tRNA-syntHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 430 AA; 47985 MW; 2E3FE99BC1E9979A CRC64;

Query Match 61.5%; Score 1382; DB 5; Length 430;
 Best Local Similarity 61.6%; Pred. No. 5.7e-113;
 Matches 263; Conservative 63; Mismatches 83; Indels 16; Gaps 3;

QY 10 DYKADCCPGNPAFNSNGPDA-----TEAE-----EDFYDPTVOTSSAKGIDYD 54
 DB 3 DTKETIVEGVALTLNKGKPDAPETVETGDAQAQAGATAPEDVDVDPNNVASSNDAGDYD 62
 QY 55 KLIVFGSSKIDKELINRIERATGQRPHPFLRGIFFSHRDMNOVLDAYENKKPFYLYTG 114
 DB 63 KLIRFGSSKIDELIARFEKITGKPAHHFIRGMPFSHDLHTLTLRQGKPFYLYTG 122

DR MGD; MGI:104630; WARS.
 DR InterPro: IPR001412; tRNA-synL_I.
 DR InterPro: IPR002306; tRNA-synL_trp.
 DR InterPro: IPR000738; WHEP-TRS.
 DR Pfam: PF00458; WHEP-TRS.1.
 DR PRINTS: PR01039; TRNASYNTHRP.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_1; 1.
 DR SEQUENCE 475 AA; 53641 MM; C3467FE8521DE4C CRC64;

Query Match 91.3%; Score 2051; DB 11; Length 475;
 Best Local Similarity 90.3%; Pred. No. 1.1e-171;
 Matches 383; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSYKAAGEDYKADCPNPAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYKLVRF 60
 DB 52 MSYKAAMEEYKACGPPGNPAGRCDSADATKASEDFVDPMTVTRSSAKGIDYKLVQF 111
 QY 61 GSSKIDKELINIRERATGQRPHRLRGIFPSHRDMNOVLDAVENKKPFYLYTGRGSSSE 120
 DB 112 GSSKIDKELINIRERATGQRPHRLRGIFPSHRDMNOVLDAVENKKPFYLYTGRGSSSE 171
 QY 121 AMHVGHLIPFTFKWLODVFNPLVIOMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 180
 DB 172 AMHGLHLPFTFKWLODVFNPLVIOMSDDEKYLMDLTLEQAYSYTVENAKDIIACGF 231
 QY 181 DINKTFIFSDLDYGMSSGFFKNVYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAA 240
 DB 232 DINKTFIFSDLEYGQSGFFRYNRYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAA 291
 QY 241 SFSNSFPQIFRDRTDIOCLIPCAIDODPFYFMTDVAFRIGYPRPALLHSTFFPALQAA 300
 DB 292 SFSNSFPKIFRDRTDIOCLIPCAIDODPFYFMTDVAFRIGYPRPALLHSTFFPALQAA 351
 QY 301 TKMSADPNSSIFLDTAKQIKTKYVKNKHAFGSGRDTIEHRQFGNCNDVDSFMYLTFFL 360
 DB 352 TKMSADPNSSIFLDTAKQIKTKYVKNKHAFGSGRDTIEHRQFGNCNDVDSFMYLTFFL 411
 QY 361 EDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEQARKREYTDIEVKEFMTPRKLS 420
 DB 412 EDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEQARKREYTDIEVKEFMTPRKLS 471
 QY 421 FDFQ 424
 DB 472 FHFQ 475

RESULT 2

Q99J58 PRELIMINARY; PRT; 481 AA.

AC Q99J58;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE TRYPTOPHANYL-TRNA SYNTHETASE.
 GN WARS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY TUMOR. WAP-TAG MODEL. 5 MONTHS OLD, GROSS TISSUE.;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003450; AAH03450.1; -.
 DR MGD; MGI:104630; WARS.
 DR InterPro: IPR002305; tRNA-synL_Ib.
 DR InterPro: IPR001412; tRNA-synL_I.
 DR InterPro: IPR002306; tRNA-synL_trp.
 DR InterPro: IPR000738; WHEP-TRS.
 DR Pfam; PF00579; tRNA-synL_Ib; I.
 DR Pfam; PF00458; WHEP-TRS.1.
 DR PRINTS; PR01039; TRNASYNTHRP.

DR PROSITE: PS00178; AA-TRNA_LIGASE_1; 1.
 DR PROSITE: PS00162; WHEP-TRS.1.
 KW Aminoacyl-tRNA synthetase.
 SO SEQUENCE 481 AA; 54325 MM; A754E1DDF58E2ER3 CRC64;

Query Match 91.1%; Score 2047; DB 11; Length 481;
 Best Local Similarity 90.1%; Pred. No. 2.5e-171;
 Matches 382; Conservative 22; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSYKAAGEDYKADCPNPAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYKLVRF 60
 DB 52 VSYKAAMEEYKACGPPGNPAGRCDSADATKASEDFVDPMTVTRSSAKGIDYKLVQF 111
 QY 61 GSSKIDKELINIRERATGQRPHRLRGIFPSHRDMNOVLDAVENKKPFYLYTGRGSSSE 120
 DB 112 GSSKIDKELINIRERATGQRPHRLRGIFPSHRDMNOVLDAVENKKPFYLYTGRGSSSE 171
 QY 121 AMHVGHLIPFTFKWLODVFNPLVIOMTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 180
 DB 172 AMHGLHLPFTFKWLODVFNPLVIOMSDDEKYLMDLTLEQAYSYTVENAKDIIACGF 231
 QY 181 DINKTFIFSDLDYGMSSGFFKNVYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAA 240
 DB 232 DINKTFIFSDLEYGQSGFFRYNRYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAA 291
 QY 241 SFSNSFPQIFRDRTDIOCLIPCAIDODPFYFMTDVAFRIGYPRPALLHSTFFPALQAA 300
 DB 292 SFSNSFPKIFRDRTDIOCLIPCAIDODPFYFMTDVAFRIGYPRPALLHSTFFPALQAA 351
 QY 301 TKMSADPNSSIFLDTAKQIKTKYVKNKHAFGSGRDTIEHRQFGNCNDVDSFMYLTFFL 360
 DB 352 TKMSADPNSSIFLDTAKQIKTKYVKNKHAFGSGRDTIEHRQFGNCNDVDSFMYLTFFL 411
 QY 361 EDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEQARKREYTDIEVKEFMTPRKLS 420
 DB 412 EDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEQARKREYTDIEVKEFMTPRKLS 471
 QY 421 FDFQ 424
 DB 472 FHFQ 475

RESULT 3

Q9D8R9 PRELIMINARY; PRT; 329 AA.

AC Q9D8R9;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE TRYPTOPHANYL-TRNA SYNTHETASE.
 GN WARS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Araiawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Huie D.A., Kamiya K., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 22.997 Seconds
(without alignments)
3189.442 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471
Perfect score: 2246
Sequence: 1 MSYKAAGGEDYKADCPGPNP.....VTDEIVKEFWTPKRLSDFQ 424

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2051	91.3	475	11 Q9DC65	Q9DC65 mus musculus
2	2047	91.1	481	11 Q9J058	Q9J058 mus musculus
3	1629	72.5	329	11 Q9DBR9	Q9DBR9 mus musculus
4	1537	68.4	305	11 Q70184	Q70184 cavia porce
5	1382	61.5	430	5 Q9U4Y1	Q9U4Y1 drosophila
6	1379.5	61.4	420	5 Q9U4Y0	Q9U4Y0 drosophila
7	1378	61.4	430	5 Q9V8G2	Q9V8G2 drosophila
8	1301	57.9	402	10 Q9SR15	Q9SR15 arabidopsis
9	1027	45.7	324	5 Q9U1R2	Q9U1R2 caenorhabdi
10	910.5	40.4	406	17 Q976M1	Q976M1 sulfolobus
11	907	40.5	386	17 Q97ZX0	Q97ZX0 sulfolobus
12	800.5	35.6	490	5 Q9U1F5	Q9U1F5 leishmania
13	682	30.4	136	6 Q9S2S5	Q9S2S5 sus scrofa
14	626.5	27.9	301	17 Q9S9S4	Q9S9S4 pyrococcus
15	534.5	23.8	380	17 Q9HN66	Q9HN66 halobacteri
16	388.5	17.3	136	6 Q9TS88	Q9TS88 bos taurus

17	310	13.8	111	5 Q95YL8	Q95YL8 encephalito
18	299.5	13.3	426	17 Q978Y8	Q978Y8 thermoplasma
19	294	13.1	157	5 Q9U5J3	Q9U5J3 trypanosoma
20	286	12.7	109	4 Q9UD15	Q9UD15 homo sapien
21	274.5	12.2	513	17 Q9HN83	Q9HN83 halobacteri
22	268.5	12.0	426	17 Q9H1W5	Q9H1W5 thermoplasma
23	188	8.4	364	17 Q9YA64	Q9YA64 aeropyrum p
24	185.5	8.3	341	16 Q97N42	Q97N42 streptococc
25	182.5	8.1	340	16 Q9YX44	Q9YX44 streptococc
26	181.5	8.1	331	17 Q9Y9Z1	Q9Y9Z1 thermoplasma
27	174.5	7.8	341	16 Q9C0D1	Q9C0D1 lactococcus
28	167	7.4	351	16 Q9RVD6	Q9RVD6 delnrococcus
29	164.5	7.3	895	10 Q9SGN2	Q9SGN2 arabidopsis
30	163	7.3	460	10 P93018	P93018 arabidopsis
31	158.5	7.1	375	17 Q9V0Z7	Q9V0Z7 pyrococcus
32	154.5	6.9	102	1 Q07119	Q07119 halobacteri
33	152.5	6.8	682	5 Q9N9B8	Q9N9B8 leishmania
34	149.5	6.7	327	17 Q9HN62	Q9HN62 halobacteri
35	149.5	6.7	408	10 P93363	P93363 nicotiana t
36	146	6.5	375	17 Q58739	Q58739 pyrococcus
37	145	6.5	332	17 Q9HKT3	Q9HKT3 thermoplasma
38	142.5	6.3	528	11 Q91WQ3	Q91WQ3 mus musculu
39	139	6.2	294	17 Q96TV3	Q96TV3 sulfolobus
40	135	6.0	365	16 Q9KNV7	Q9KNV7 vibrio chol
41	134.5	6.0	525	5 Q9VW60	Q9VW60 drosophila
42	130.5	5.8	419	16 Q9ZBB1	Q9ZBB1 listeria in
43	129	5.7	339	2 Q9KZK7	Q9KZK7 streptomyce
44	124	5.5	344	16 Q9AC05	Q9AC05 caulobacter
45	123	5.5	347	10 Q82313	Q82313 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q9DC65	PRELIMINARY;	PRT;	475 AA.
AC	Q9DC65;	01-JUN-2001 (TREMBLREL, 17, Created)			
DT	01-JUN-2001 (TREMBLREL, 17, Last sequence update)				
DT	01-DEC-2001 (TREMBLREL, 19, Last annotation update)				
DE	ADULT MALE LUNG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,				
DE	CLONE:1200002C07, FULL INSERT SEQUENCE.				
GN	WARS.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RC	SEQUENCE FROM N. A.				
RP	SPRAIN=C57BL/6J; TISSUE=LUNG;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanao I.,				
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,				
RA	Sakurai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaide M.F.,				
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombauts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,				
RA	Hayashizaki Y.,				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK004541; BAB23357.1; -				

Query Match	7.78;	Score 172;	DB 1;	Length 394;
Best Local Similarity	22.78;	Pred. NO. 7.8e-07;		
Matches	88;	Conservative	62;	Mismatches 132;
			Indels	106;
			Gaps	22;

```

QY 94 RDMNQVL-----DAYE-NKKPFYLTCRGSSSEAMVGHILPFIETKMLQDFENV--PL 144
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 KNLOEVLNPOILKIVLEVQKRHLKLYMGTAFTGPR-HCGTFVP--MKR-LADELAKAGEV 73
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 VIQMTDDEKYL-----WKDITLD--QAYSDAVENAKDIIAGCFDINKTF 186
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 TVLLADILHAFIDNNKAPLEVVNRYAKYETTELKALISINWPIEKLTFVGVSSYQLTDPV 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 IFSDDLDMGMSGGYKKNVVK-----LOKHVFNQVKIFEGFTSDCIKGISFPAIQAPS 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 ---TMDIFRLSNIVSQDARKAGADVQYANPLLSGLI-----YPLMOQA-- 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 FSNSEFPOLFRDRTIOCLIPCAIDQDDEYFMRTRIVARIGPKPALHSTFEFPAL--OGAQ 300
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 -----LDEQFLDVCDFG-GVUQKRIFLVAEENLPSLSIGKKRAHLMNPMVFGLAOGC- 226
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 TKMSASDPNNSIFLDTAKOIKTKVNKHAESGGRDITEH-----ROFGS 345
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 -KMSASDPNNSKIDLEBPQKXKXINSAPFCSPG--NVEENGLLSFVQYIAPIQELFCGT 283
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 346 NCDYDVSEFMILTFFLEDDDK-----LEQIRKDYTSCAMLTGELK---KALIEVLO 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 N-----HEEPIIDRPEKFGGPIITYKSEEMKLAKEERKLSPDPDLKIGVADAINELLE 355
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 393 PLIAEHQARREVTDELVEYKF--WTPRK 418
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 P-IOEFRANKEFOEASEKGYPAITPOK 362
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: October 24, 2002, 12:51:43
Job time : 8.77387 secs

```

AC Q46127;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
GN TRPS OR TRSA.
OS Clostridium longisporum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1523;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6405;
RA Brown G.D., Thomson J.A.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC -i- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -i- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -i- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L49336; AAC05711.1; -.
DR HSSP; P00953; 1D2R.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTTRP.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 12 20 "HIGH" REGION.
FT SITE 201 205 "KMSKS" REGION.
FT BINDING 204 204 ATP (BY SIMILARITY).
SO SEQUENCE 341 AA; 38256 MW; 692C820F5A08E4D1 CRC64;

Query Match 7.8%; Score 175.5; DB 1; Length 341;
Best Local Similarity 24.7%; Pred. No. 3; 3e-07;
Matches 78; Conservative 53; Mismatches 150; Indels 35; Gaps 11;

OY 111 LVTGRSPSEAMVGHLPFIPTK-WLQDVFNVPVLTQMTDDEKYLIMKDLTLDQAYGDVAV 169
DB 6 ILTGDRPTGK-LHIGHVSLKNRYQLQNSGDRSPIMADQALTDNARNPEKIRNSLI 64
OY 170 ENAKRIIACGFINKTFIF--SDLDYMGSSGFYKNV---KIQCHVFNQYKGIKFGPLD 224
DB 65 EVALIDYLAAGIDPLKSTLIYQSQIPELNELTMHYLNLVTLNLENNPNYKAIKQKNFEN 124
OY 225 SDICIKISPPAIQAQSPFSNPQIFRDRDIOCLIPCAIDDDPFYRMTRDVAPRIG--- 281
DB 125 STIPAGFLIYPSQADITR-----FKATP-----VPVGEDDLPIMIEQARELYVRSFNIT 173
OY 282 -----YKPRALHSTF--FPALQAGQTKMASADPNSSIFLTDTKAKQIKTKVKNKHAFFSG 333
DB 174 GKEVLVERKAVIPKGTIGRLPCTDG-KAKMSKSGN-AIYLADEADVIKQVMSMTDPN 231
OY 334 RTIEHNRFGGNCVDVSMFLTFLEDDDKLEQIRKDYTGAMITGLKALLLEVLP 393
DB 232 KIKVDPDGVEN---TVETYLDFCKDTETLEMKAHVSGGLGVKKVFLNELIQA 287
OY 394 LIAEQARKREVTDEI 409
DB 288 ELEPIRNKRKEFOKDI 303

```

```

RESULT 15
SYSC_YEAST
ID SYSC_YEAST STANDARD: PRT: 394 AA.
AC P36421;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tyrosyl-tRNA synthetase, cytoplasmic (EC 6.1.1.1) (Tyrosyl--tRNA
DE ligase) (TYRS)
GN TyS1 OR MGM104 OR YGR185C OR G7522.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93286133; PubMed=8509419;
RA Chow C.M., Rajbhandary U.L.;
RT "Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene.
RT Isolation by complementation of a mutant Escherichia coli suppressor
RT tRNA defective in aminoacylation and sequence analysis.";
RL J Biol. Chem. 268:12855-12863(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Guan M.-X., Chen X.-J., Clark-Walker G.D.;
RL Submitted (MAY-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363(1997).
RN [4]
RP SEQUENCE OF 1-36 FROM N.A.
RC STRAIN=Bj926;
RX MEDLINE=95087887; PubMed=7995524;
RA Henry N.L., Campbell A.M., Fearer W.J., Poon D., Weil P.A.,
RA Kornberg R.D.;
RT "PEIF-TAF-RNA polymerase II connection.";
RL Genes Dev. 8:2868-2878(1994).
CC -i- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -i- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L12221; AAB59329.1; -.
DR EMBL; X71998; -. NOT ANNOTATED_CDS.
DR EMBL; Z72870; CAAG7211.1; -.
DR EMBL; X99074; CAAG7529.1; -.
DR EMBL; U13015; AAG61641.1; -.
DR PIR; A45999; A45999.
DR SGD; S0003417; TYSL.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; tRNA-synt_tyr.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTTHYR.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 48 56 "HIGH" REGION.
FT SITE 227 231 "KMSKS" REGION.
SO SEQUENCE 394 AA; 44020 MW; 57EBDB9BE6D054B7 CRC64;

```

```
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000664; BAA81476.1; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 81 89 "HIGH" REGION.
FT SITE 258 262 "KMSKS" REGION.
SQ SEQUENCE 374 AA; 42400 MW; A72635B7CA3F9189 CRC64;

Query Match 15.7%; Score 353; DB 1; Length 374;
Best Local Similarity 30.4%; Pred. No. 6.6e-22;
Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 38 VDPWTVQVSSAKIDYDKLYRFGSSKIDKELINIERATGGRPHNPLRGIFFSHRDM 97
DB 8 LDPW---GAVEIKDYDLRLTFTGIRPSEVL--PLLRKAGSEPSFLMRKGIIFGHRDE 61
QY 98 OVLDAENKRPFLYLTGRGPSEAMHVGHLIPFTKMLQ-DVFNPVLIOWTDEKILM 156
DB 62 KILEKAKGERAVALTGTMPGSK-FHFGHKLVLDLILOKNGFV--FVALADEAAV 118
QY 157 KDLTLDOAGDAVEN-ANDIACGFINKT-FIFSDLYMGMSGFYKNVAKIOKHVEN 214
DB 119 RRIREEAVRIAEVEYIANNMIALGLDPKDTREYFQ---TNGPTPYRLIQFSGKVTA 174
QY 215 OVKGIFG-FTPSDCIGKISFPAIOAAPSNSFPQIFRDRDIOCLICADDODPYFMT 273
DB 175 EMEALYGLTTPAKKMASLT---QAADILHVQLDDEYGYR--HVVVGVGADQDPHRLT 227
QY 274 RDVAPR---IGYRPALLHSTFFPALOGAQTKMSADPNSSIFLTPRAKQIKTKVNKA 329
DB 228 RDLADRMAGVVELERPATYHKLPGLDG--RKMSSTRDSTIFLTPREYAKNKLFR-A 284
QY 330 FSGGHDITIEHRQFGGNCV-DVSEFMYLTFLLEDKLEQIRKDYTS---GAMLTGELK 385
DB 285 LTGGRATAEQRRILGVEPVCSVYHMDYLHMPRDGGEVKNH---YTSCLRLKILGCECK 341
QY 386 ALIEVLOPLIAHQARRKEVDELYKEFMTPR 417
DB 342 IAMEKLERFLAEHOSRLEKAKTIAMKLVPPER 373

RESULT 13
SYN_ARCFU STANDARD; PRT; 323 AA.
AC 029482;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tyrosyl-tRNA synthetase (Ec 6.1.1.1) (Tyrosine--tRNA ligase) (TYRS).
GN TYRS OR AF0776.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kechum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spilgys T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001051; AAB90462.1; -
DR TIGR; AF0776; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; tRNA-synt_tyr.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 41 49 "HIGH" REGION.
FT SITE 214 218 "KMSKS" REGION.
SQ SEQUENCE 323 AA; 36616 MW; A655ABE4A5116642 CRC64;

Query Match 8.5%; Score 192; DB 1; Length 323;
Best Local Similarity 23.1%; Pred. No. 1.3e-08;
Matches 78; Conservative 67; Mismatches 127; Indels 66; Gaps 15;

QY 92 SHRMNOVLDAENKRPFLYLTGRGPSEAMHVGHLIPFTKMLQDVFNPLVIOQMTDD 151
DB 19 TEEELRQLETKRPR--AYGYEPGE-IHLGHMTVQRLMDQEA-GFEIIVLADI 73
QY 152 EKYLMKDLTLDOAGDAVENAKDIACGFDINKTFTFSDLYMGMSGFYKNVAKIOKHV 211
DB 74 HAYLNEKTFEEIEAVADYNNKVFALGLDSRAKFLVSGEYQ-LSRIYVDVLMKMATIT 132
QY 212 TFNOYK---GIFGTFSDCIGKISFPAIOAAPSNSFPQIFRDRDIOCL-IPCA--- 263
DB 133 TLNRARRSMDEVSRKKEPDMVSQMIYPLMQA-----LDIAHLGVDLAVGG 177
QY 264 IDQDPYFMTRDVAPRTGYPPKALLHSTFFPALOGAQTKMSADPNSSIFLTPRAKQIKT 323
DB 178 IDORKIHMLARENLPRLGYSSPVCLHPILVGLDG--OKMSSSKGN-YISVRDPPEVER 234
QY 324 KVNK-HAFSG-----GRDTIEHRQFGGNCV-DVSEFMYLTFLLEDKLEQIRKDYTS 363
DB 235 KIRKAYCPAGVAEENPILDIKAYHILPREGKIIVERDAKFGG--DVE---YASF----- 283
QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQARR 401
DB 284 ---EELADEFSQGLHPDLTKIAVAKYLMMLLEDARRK 318

RESULT 14
SYN_CLOLO STANDARD; PRT; 341 AA.
ID SYN_CLOLO
```



```

FT NON_TREE      134      134
SQ SEQUENCE      134 AA; 15744 MW; 60E2935B7E1E344F CQC64;

Query Match          17.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 2.9e-25;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

OY    35 EDFVPMWTVQTSAR-----GIDYDKLYIFRFGSKIDKELINIERATGOORPHFRRGTF 90
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB    3  EQRITPMDVEVSTDEVPAYDAIDYKDITNIOFGCEKNQALADLELKSCKPAHYFFRRGV 62
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY    91 FSHRDMMNOVDAYENKKPFYLYTGRGPSSSEAMVGHLPFIETFKWLQDVFNPLYIQMTD 150
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB    63 FAHRPFINLLDIANNRPFYLYTGGRGPSKTMHIGHTIPIFLCLKRYMQDAFKIRLVIGLT 122
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY    151 DEKYLMKDLTTD 162
       |||:||||:|:|:
DB    123 DEKLWKSMRLTE 134

RESULT 11
SYM_ARCFU
ID     SYM_ARCFU      STANDARD:      PRT:      420 AA.
AC     O28579;
DT     15-JUL-1998 (Rel. 36, Created)
DT     15-JUL-1998 (Rel. 36, Last sequence update)
DT     16-OCT-2001 (Rel. 40, Last annotation update)
DE     Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE     (TrpRS).
GN     TRPS OR AF1694.
OS     Archaeoglobus fulgidus.
OC     Archaeae; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC     Archaeoglobus.
OX     NCBI_TaxID=2234;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX     MEDLINE=98049343; PubMed=9389475;
RA     Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA     Kechum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA     Richardson D.L., Kerlavage A.R., Graham D.E., Kyriides N.C.,
RA     Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA     Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA     Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodok A., Zhou L.,
RA     Colton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA     Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA     Mason J.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA     Venter J.C.;
RT     "The complete genome sequence of the hyperthermophilic sulphate-
RT     reducing archaeon Archaeoglobus fulgidus.";
RL     Nature 390:364-370(1997).
CC     -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC         diphosphate + L-tryptophanyl-tRNA(Trp).
CC     -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC     -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC     or send an email to license@sib-sib.ch).
CC
CC     -----
CC     EMBL; AE000986; AAB89554.1; .
DR     TIGR; AF1694; .
DR     InterPro; IPR002305; tRNA-synt_1b.
DR     InterPro; IPR001412; tRNA-synt_1.
DR     InterPro; IPR002306; tRNA-synt_ttp.
DR     Pfam; PF00579; tRNA-synt_1b; 1.
DR     PRINTS; PRO1039; TRNASYNTHTRP.
DR     PRSITE; PS00178; AA_TRNA_LIGASE_I; FALSE NEG.
```

KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome. "HIGH" REGION.
 FT SITE 72 80
 FT SITE 308 312 "KMSK" REGION.
 SQ SEQUENCE 420 AA; 47283 MM; 9315152EF172F24 CRC64;

 Query Match 16.5%; Score 370.5; DB 1; Length 420;
 Best local similarity 27.3%; Pred. No.2.7e-23;
 Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

 QY 38 VDPWTVGVTSSAKGIDYDKLIVRFGSSKIDKLEINRIERATGGRPHNFKRGIFGSHRDMN 97
 Db 3 VYPMEVEGV---IDYSKLIIEFGMQPF-SEVLPEDID-----NPHILMRGALFGRDWM 52
 QY 98 QVLDVAENKKRPFLYLTGRKRPSSAMHVGHLFFITKWLQDVFNVPVLYQMDDDEKYLAK 157
 Db 53 RLTEMOKKEPRVAVVSGFMRPG-LRPHGKMKMDELVWQASGAKFV-ALADMEAHSVR 110
 QY 158 DLTLDQAVGDAVENKAKDIATACGDFINKTFFESDLDYMGSGSGFYKNV-KIQNHVFNQV 216
 Db 111 GLSWKTRRELGLMYIKSIIALDRDAVIFYGSS-----KSHYVADLAEFLSAEVENFSL 164
 QY 217 KQIFGFTSDICIGKISFPAQAPRSPNSFOIFDRDRIDQCLICALDDQRYFMTRDV 276
 Db 165 RAIFYGNSDTSIAKMEVTAIQADIL---HPQLSDFGGKRPVAVVGVADODPHMRILTRDL 221
 QY 277 ARI----- 280
 Db 222 AARISIFSEPRPEGVRVRSKGAELYSLRLPREDKTYEEHMDIFGEAEITEAVRKI 281
 QY 281 -----GYPRALLHSTFFPALQGAQTMSASDPNSIFLIDTAKIKTKVKNKHAFFSG 333
 Db 282 EVEIGFAPRIPSSYTHRETTGLTG--KMSSEKESYISLIDPREEGAKVMK-AFTGG 338
 QY 334 RPTIEHQFGNCQVDVSFMVLTFFLED-DOKLQIKRDYTSAGMLGELKALLLEVLYQ 392
 Db 339 RAFAEQRRLGGEPRCVFELLYSPHLIDSDLEWQIEACREGRLCGCKKMAAEIVK 398
 QY 393 PLIAEQARKEV 405
 Db 399 SFLKEHQEKMEAV 411

 RESULT 12
 SYM_AERPE STANDARD; PRT; 374 AA.
 AC OY924;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE (TrpRS).
 DE tryptophanyl-tRNA synthetase (EC 6.1.1.3) (tryptophan--tRNA ligase)
 GN TRPS OR APE2461.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 CC Aeropyrum.
 CX NCBI_TaxID=56636;
 RN NCB1
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jinno K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takemaki M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
 RA Yamazaki J., Kashiida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.

Db 113 RNMSPTTKELALNEYINIALGLDPEKINVIYLSQKYQV-----KDLALILSKRTWMS 167
QY 215 QVKGIFGTDSDCIGKISFPALQAPSFNSFPQIFRDRTP--DIOCLIPCAIDDPPYRM 272
Db 168 EMKAIYGFKEGEMTNGHVPAPIVQADIL--HPOLDENI,SPBPKVVPVGVGIDDPHRL 224
QY 273 TRDVAAPR---IGYPRPALHSTFPALQAGTQKMSASDPNSIFLTPAKOIKTKTVKNKA 329
Db 225 TRDIANRAKKERKFTPPSSTYHFMGTGLGG--KMSSKPEFAIFLTDDEKTVKKKIFS-A 281
QY 330 FSGGRTDIEEHRQFG---NCDVDVSFMYLTFEEDDKLPQIRKDYTSGAMLTGELKKA 386
Db 282 KTGGRTELLEHKYKGVGEVVECYVELFLY--HLIUDKELAEIYQKCSGELTGCKCKM 339
QY 387 LIEVLIQPLIAEHQARKKEVTEIYK 411
Db 340 AYERVVEFLKDKERQAKELAVK 364

RESULT 9
SYM_METH STANDARD: PRT: 364 AA.
ID SYM_METH
AC 026352;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TPRS).
GN TRPS OR MTH251.
OS Methanobacterium thermoautotrophicum.
OC Archaea: Euryarchaeota: Methanobacteriales: Methanobacteriaceae:
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Potlter B., Qiu D.,
RA Spadatore R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McQuail S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-Tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000812; AAB84757.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PRO1039; TRNASYNTTRP.
DR PROSITE: PS00178; AA tRNA LIGASE. I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
KW Complete proteome.
FT SITE 70 78 "HIGH" REGION.
FT SITE 251 255 "KMSK" REGION.
SQ SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

Query Match 17.7%; Score 397.5; DB 1; Length 364;
Best Local Similarity 27.9%; Pred. No. 1.3e-25;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;
QY 38 VDPWTVOTSSAKGIDYDKLIVFGSSKIDKELINIKERATGQRPHEFLRGIFFSHRDM 97
Db 2 IDPW-----GSAK-LEYQDLINFGVRP-SEVLEDEV-----PEFSWLMRRGILIGHROYE 50
QY 98 QVLDAYENKKFFPYLYTGSGPSEAMHGHILPFTKYLQVFNVPYLYQNTDDEKILMK 157
Db 51 RIISAMKKGEDEFAVATGMPGSR-MHIGKMIVDLRV-YRMAEAIFIPADMEAVSAR 108
QY 158 DLTLDQAGDGVEN-ANKDIACGFDYK-----TFJESDLDYWGMSGFFKYNVYK 206
Db 109 GVDFEDSRRRIAEIYIAGYIALGLDLEKDNHLYLQSNLMVEDIAY----- 156
QY 207 IQKHVFNQVKGIFGTDSDCIGKISFPALQAPSFNSFPQIFRDRTPDIOCLIPCAIDQ 266
Db 157 LAGKVNFEELRAIYGFSTGSTMAMHYAPIIOVSDILHPQDELGPR---PVIYVGPDQ 213
QY 267 DPYRMTDVAAPRI---GYPRPALHSTFPALQAGTQKMSASDPNSIFLTPAKOIK 322
Db 214 DPHRLTRDIAARFRDRYGFILPSTYHFMGTGLGG--KMSSNRKPSAIFLSDTPEAE 271
QY 323 TKVKNHAFSGGRDIEEHRQFGNCDVDVSFMYLTFEEDDKLPQIRKDYTSGAMLTG 381
Db 272 AKI-RNATGREGRELKKERELGVPEECITETLLIHMSGSDSLIEETECRNGTLMCG 330
QY 382 ELKRALIEVLQPLIAEHQARKKE 404
Db 331 ECKNNTAEFIKFEELSVKREK 353

RESULT 10
SYM_ENCCU STANDARD: PRT: 134 AA.
ID SYM_ENCCU
AC 096771;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TPRS) (Fragment).
OS Enecephalitozoon cuniculi.
OC Eukaryota: Microsporidia; Unikaryoniidae; Enecephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98277683; PubMed=9615449;
RA Peyretailade E., Broussoille V., Peyret P., Metenier G., Gouy M.,
RA Vlavres C.P.;
RT "Microsporidia, amitochondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-Tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ012470; CAA10034.1; -
DR HSP: P00952; IMYC.
DR InterPro: IPR001412; tRNA-synt_1.
DR PROSITE: PS00178; AA tRNA LIGASE. I; 1.
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
FT SITE 89 98 "HIGH" REGION.

```

SYW_PYRAB
ID SYW_PYRAB STANDARD: PRT: 385 AA.
AC 09UT11:
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR PAB1111.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_taxid=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
di-phosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ248288; CAB5601.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTTRP.
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
KW Complete proteome.
FT SITE 82 "HIGH" REGION.
FT SITE 253 "KMSKS" REGION.
SQ SEQUENCE 385 AA; 45100 MW; 4C29D01414976B12 CRC64;

Query Match 35.8%; Score 803; DB 1; Length 385;
Best Local Similarity 45.3%; Pred. No. 2.8e-59;
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 35 EDF-VDPWTVOTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQRPHHFLRGIFFSH 93
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 3 EDFKVTPEVEGV---VDYNKLIEHFGTSPLEELKELTALTELSPLFFRRKFFESH 58
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 94 RDMNOVLDAVENKKPFYLYTGRGSSSEAMHGHILPIFTKWLQDVNPVLYIQMTDDEK 153
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 59 RDYDVLQDYEGRGFFLYTGRGPG-PMHIGHILPIFPATKWLQKFGVNIYLIQTDEK 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 154 YLMKD-LTLDAQYDAVNNAKDIAGCGDINKTFESDLDMGMSGFGYKNVVKQKQIVT 212
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 FLEKNTLFEDTKHAYENLIDIIAVGPDPTFTFQNSEF---TKLYEMALPIAKKIN 173
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 213 FNOVGIGFTLSDICIGKISFPAIDAAFSFNSFQIFRDRITDIOCLIPCAIDDDPYFRM 272
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 174 FSMARAAVGFTEFQSKIMFFPAIDAIPTF-----FEKR---RCLIPALADDDPYRKL 223
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 273 TRDVAIRIGYKPPALLHSTFEPALOGAOTKMSASPNSSIFLTDTAKOIKTKVKNHAFSG 332
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 224 QRDFAESLGYYTAIHSKFVPSLTSLSGKMSASKEPTAIYLTDSPEDEVKKVMWFALTG 283
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 333 GRDTIEEHROPGNCDDVYSEFWYLFLEDDDKLEQIRKDY---TSGAMLTGELKKAILE 389
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 284 GRPTLEOREKRGPEKCEKVFEMLEIFFEEDK--KLKERIYACKNGELTGCECKRYLIS 341
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 390 VLOPLIAHQARKEVTEIYK 411

```

```

Db 342 KIOEFLKEHQKRRKKAKEQIEK 363
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
RESULT 8
SYW_METJA
ID SYW_METJA STANDARD: PRT: 370 AA.
AC 058810:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_taxid=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; Pubmed=6688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Colton M.D., Roberts C.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
di-phosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67582; AAB99425.1; -
DR TIGR: MJ1415; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
KW Complete proteome.
FT SITE 75 "HIGH" REGION.
FT SITE 255 "KMSKS" REGION.
SQ SEQUENCE 370 AA; 42660 MW; E6C71107CF82B59D CRC64;

Query Match 18.2%; Score 409.5; DB 1; Length 370;
Best Local Similarity 30.9%; Pred. No. 1.3e-26;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

QY 40 PWTVOTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNV 99
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 8 PW-ETPAV--IDYKKTMEQFGVKPIVDVLDLKEE-----HHFFRRNIIILGHRDFERI 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 100 LDAYNKKPFYLYTGRGSSSEAMHGHILPIFTKWLQ---DVFNVPVLYIQMTDDEKILM 156
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 VDAIKNNKEFNAVSGMPSGK-MHFGHRAVVDLKEFYOKYTDNINIPt---ADEAEAWA 112
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 157 KDLTLDAQYDAV-ENAKDIAGCGDINKTFESDLDMGMSGFGYKNVVKI-QKHVTFN 214

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50142; CA90500.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1rp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; tRNA-synt_1rp.
DR PROSITE; PS00178; AA-TRNA-LIGASE.1; 1.
KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;
KW Ligase; ATP-binding.
FT SITE 91 100 "HIGH" REGION.
FT SITE 275 279 "KMSK" REGION.
SQ SEQUENCE 395 AA; 44910 MW; E656AEB876C5FD9 CRC64;
Query Match 53.9%; Score 1210; DB 1; Length 395;
Best Local Similarity 59.6%; Pred. No. 4.3e-93;
Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;
QY 34 EDEVDPTVOTV-----SAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHELRGI 89
DB 4 EEQIYTPMDVAGSIVDGEKIDYERLIVQGTGRTPEQLEFEREKLGKPKHLLRRGA 63
QY 90 FFHSDMNQVLDAYENKRPFLYLTYGRGSSSEMAHGHILPFIETKWLQDVFNPLVDMT 149
DB 64 FFHSDPFIIMIDRYEQKKPFLYLTYGRGSSSMHGHMIPFECMKWLQDVFNPLVDMT 123
QY 150 DDEKTLKMD-LTLQAVYADAVENAKDIIACGFDINKTFIFSLDYMGSNGFYKVVVQ 208
DB 124 DDEKFLFQGVSLDCCQAFARENAKDIIAVGFDPKTITFENSTYVG--GAFYQVAVR 181
QY 209 KHVTFENYKGIFFETSDICIGISFPALQAAAPFSNSPQIFRRTDIOCLIPCAIDDP 268
DB 182 KCIANQSKACGCFDSDISGKIHFAISIQAAAPFSNSPPIFNAGAKDIPCLIPCAIDDP 241
QY 269 YFRMTQVAPRIGYPRKPLHSTFPALQAGACTKMSADPNSSITFLTAKQIKTKVNH 328
DB 242 YFRMTQVAPRIGYPRKPLHSTFPALQAGACTKMSADPNSSITFLTAKQIKTKVNH 301
QY 329 AFSGGRTIEEHRQFGNCVDVSPMYLTFELEDKLEQIRKDYTSAGMLTGEKALKI 388
DB 302 AFSGGGATIEIHRKGGKPPDVAVAYQYLSFLDDEKILQKLYNTYKAGTSLGEMKGECT 361
QY 389 EVLQPLIAEHQARKEVTDEIVKEFT--PKRLSF 421
DB 362 KLLQGFVSDPQARSKVDEATLDMFMDSRKLEW 395
RESULT 6
SYMC_YEAST STANDARD; PRT; 432 AA.
ID SYMC_YEAST
AC 012109;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (tryptophan--
DE tRNA ligase) (Tfprs).
GN WRS1 OR YOL097C OR HRA32.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96076631; PubMed=7502582;
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
RT *Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV

RT including the Ty1-H3 retrotransposon, the sufl(+) frameshift
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Tyr-1a and a
RT delta element.";
RN Yeast 11:1069-1075(1995).
RN [2]
RP FUNCTION.
RX MEDLINE=97197969; PubMed=9046085;
RA John T.R., Ghosh M., Johnson J.D.;
RT *Identification and expression of the Saccharomyces cerevisiae
RT cytoplasmic tryptophanyl-tRNA synthetase gene.";
RL Yeast 13:37-41(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Tyr) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Tyr).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z48149; CA88164.1; -
DR EMBL; Z74839; CA99110.1; -
DR SGD; S0005457; WRS1.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; tRNA-synt_1rp.
DR PROSITE; PS00178; AA-TRNA-LIGASE.1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 111 120 "HIGH" REGION.
FT SITE 295 299 "KMSK" REGION.
SQ SEQUENCE 432 AA; 49350 MW; C408F16973769736 CRC64;
Query Match 51.8%; Score 1163; DB 1; Length 432;
Best Local Similarity 54.8%; Pred. No. 4e-89;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;
QY 30 ATEAEEDFVDMTV-----QTSASAGIDYDKLIVRGSSKIDKELINRIERATGQRPHE 84
DB 19 STDVKEQVTVTMDVAGVDEGQRAQNIIDYDKLQFGTKPVNEETLKRFGQVGREPHF 78
QY 85 LRGIFFSHRDMNOVLDAYENKRPFLYLTYGRGSSSEMAHGHILPFIETKWLQDVFNPL 144
DB 79 LRGIFFSERDFTKLLDIEGQKPFELTYGRGSSSMHGHMIPFETKWLQDEVFVPL 138
QY 145 VIQMTDDEKYLWK-DLTLQAVYADAVENAKDIIACGFDINKTFIFSDIDYMGSSGFYKN 203
DB 139 VIELTDDKFLFKHKLITINDYKNFARENAKDIIANGFPKNTFIFSDIQYMG--GAFYET 196
QY 204 VKIQRHVTENOVKGIFFETSDICIGISFPALQAAAPFSNSPQIFRRTDIOCLIPCA 263
DB 197 VVRVSQIITGSTAKAVFGFNDSDICIGKHFHFAISIQAAAPFSNSPPIFNAGAKDIPCLIPCA 256
QY 264 IDQPYFRMTQVAPRIGYPRKPLHSTFPALQAGACTKMSADPNSSITFLTAKQIKTK 323
DB 257 IDQPYFRMTQVAPRIGYPRKPLHSTFPALQAGACTKMSADPNSSITFLTAKQIKTK 316
QY 324 KVNKAFFSGRDTIEEHRQFGNCVDVSPMYLTFELEDKLEQIRKDYTSAGMLTGEK 383
DB 317 KINKAFFSGGVSAADLIELHSGNPPDVAVAYQYLSFLFKDDVFLKCYKISGELLSEW 376
QY 384 KKAILEVQLPLIAEHQARKEVTDEIVKEFTPKRL 419
DB 377 KKCIETIQEFYKAFQERRAQVDEFTLQKFMVPHKL 412
RESULT 7

```

|||||
Db 173 MHVGHILPFIETKWLQDVDPVLYVOMSDDEKYLKKDLTLEQVGYGTLENKADIACGFD 232
OY 182 INKPTIFSDLDYMGKSSGFEYKNNVKIOKHVTFNOKYKIGFPGDSDCICKISPPAIQAAP 241
Db 223 VNKTFIFSDLDYMGKSSGFEYKNNVKIOKHVTFNOKYKIGFPGDSDCICKISPPAIQAAP 292
OY 242 FSNSEFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQAQ 301
Db 293 FSNSEFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQAQ 352
OY 302 KMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 361
Db 353 KMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 412
OY 362 DDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLSF 421
Db 413 DDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLSF 472
OY 422 DEFQ 424
Db 473 HYQ 475

```

RESULT 4
SYN_MOUSE
ID SYN_MOUSE STANDARD: PRT: 481 AA.

```

AC P32921:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=95018226; PubMed=7932716;
RA Rajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
EMBRYONIC STEM CELLS.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69656; CAA49347.1; -;
DR EMBL; X69657; CAA49348.1; -;
DR PIR; S31461; S31461.
DR MGI; MGI:104630; Wars.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.

```

```

DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PS01039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPLIC 476 481 MISSING (IN MAJOR ISOFORM).
SO SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;

```

Query Match 90.5%; Score 2032; DB 1; Length 481;
Best Local Similarity 89.6%; Pred. No. 2,8e-161;
Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

```

OY 1 MSYKAGEDYKADCPENPAPTSNHGPDATBAEDVDVPTVOTSSAKGIDYKLYRF 60
Db 52 MSYKAMGEYKAGCPENPAGRNCDSDATKASDFVDPMTVTRTSKAGIDYKLYQP 111
OY 61 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGSPSE 120
Db 112 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGSPSE 171
OY 121 AMHVGHLIPFIETKWLQDVFNVPVLYIQMTDDEKYLKMDLTDQAVGDAVENAKDIACGF 180
Db 172 AMHVGHLIPFIETKWLQDVFNVPVLYIQMSDDEKYLKMDLTDQAVSYVENAKDIACGF 231
OY 181 DINKTFISDLDYMGKSSGFEYKNNVKIOKHVTFNOKYKIGFPGDSDCICKISPPAIQAAP 240
Db 232 DINKTFISDLDYMGKSSGFEYKNNVKIOKHVTFNOKYKIGFPGDSDCICKISPPAIQAAP 291
OY 241 FSNSEFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQAQ 300
Db 292 FSNSEFPQIFRDRTDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQAQ 351
OY 301 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 360
Db 352 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 411
OY 361 EDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLS 420
Db 412 EDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLS 471
OY 421 DEFQ 424
Db 472 HFQ 475

```

RESULT 5
SYN_SCHPO
ID SYN_SCHPO STANDARD: PRT: 395 AA.

```

AC Q09692:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
ligase) (TrpRS).
GN SPAC2F.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----

```

```

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53918; CAA37872.1; -.
DR EMBL: X52113; CAA36356.1; -.
DR EMBL: M74074; AAA30799.1; -.
DR PIR: A40279; YMR0.
DR PTR: S14540; S14540.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 24 69 WHEP-TRS.
FT SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
FT SITE 169 178 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT CONFLICT 17 17 L -> M (IN REF. 2).
SQ SEQUENCE 475 AA; 53729 MW; F7E513750137EB32 CRC64;

```

```

Query Match          95.0%; Score 2134.5; DB 1; Length 475;
Best Local Similarity 95.0%; Pred. NO. 8.4e-170;
Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 2 SYKAAAGEDYKADCPGNAPTSMHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRFG 61
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 54 SYKAAATEDYKADCPGNAPTSMHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRFG 113

QY 62 SSKIDKELINRIERATGQRPHFLRRCIFFSHRDMNOVLDAVENKKPFYLYTGSPSSEA 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 114 SSKIDKELVNRIERATGQRPHFLRRCIFFSHRDMNOVLDAVENKKPFYLYTGSPSSEA 173

QY 122 MHVGHLLPFTFTKWLQOVFNPLVIOGTDDEKYLKMDLTLDQAVGDAVENAKDIIACGFD 181
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 174 MHVGHLLPFTFTKWLQOVFNPLVIOGTDDEKYLKMDLTLDQAVGDAVENAKD-ITCGFD 232

QY 182 INKTFISDLDYMGMSGFEYKNVYKIOKHVTFENOVKGIFGFTSDCIGKISFPALIOAPS 241
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 233 INKTFISDLDYMGMSGFEYKNVYKIOKHVTFENOVKGIFGFTSDCIGKISFPALIOAPS 292

QY 242 FSNFSPQIFRDRDIOCLIPCALIDODPEYFRMTADVAPRIGYKPKALLHSFFPALQAGQT 301
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 293 FSNFSPQIFRDRDIOCLIPCALIDODPEYFRMTADVAPRIGYKPKALLHSFFPALQAGQT 352

QY 302 KMSASDPNNSIFLDTAKQIKTKYKNAHAFSGRPTIEHQFGNCVNDVDSFMTLTFPLE 361
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 353 KMSASDPNNSIFLDTAKQIKTKYKNAHAFSGRPTIEHQFGNCVNDVDSFMTLTFPLE 412

QY 362 DDDLEQIRKDYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSF 421
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 413 DDDLEQIRKDYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSY 472

QY 422 DFQ 424
   |||
DB 473 DFQ 475

```

RESULT 3
SYM_RABIT
ID SYM_RABIT STANDARD; PRT; 475 AA.

```

AC P23612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (trpRS).
GN WARS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RX MEDLINE=94009008; PubMed=8404867;
RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Druegon G.,
RA McCaughan K.K., Kisselev U.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins.";
RL EMBO J. 12:4013-4019(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33460; AAA31246.1; ALT_SEQ.
DR EMBL: U02595; AAB60257.1; -.
DR PIR: A35904; YMRBPR.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF43DC4 CRC64;

Query Match          91.5%; Score 2056; DB 1; Length 475;
Best Local Similarity 90.5%; Pred. NO. 2.8e-163;
Matches 383; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

QY 2 SYKAAAGEDYKADCPGNAPTSMHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRFG 61
   |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 53 SYKAAAGEDYKADCPGNAPTSMHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRFG 112

QY 62 SSKIDKELINRIERATGQRPHFLRRCIFFSHRDMNOVLDAVENKKPFYLYTGSPSSEA 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 113 SSKIDKELVNRIERATGQRPHFLRRCIFFSHRDMNOVLDAVENKKPFYLYTGSPSSEA 172

QY 122 MHVGHLLPFTFTKWLQOVFNPLVIOGTDDEKYLKMDLTLDQAVGDAVENAKDIIACGFD 181

```

RA Erolova L.Y., Grigorjeva A.Y., Sudomolina M.A., Kisselev L.L.,
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-
 RT response elements and exon-intron organization.";
 RL Gene 128:237-245(1993).
 [7]
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Raamsen H.H., van Damme J., Puype M., Geeser B., Celis J.E.,
 RT Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [8]
 RP FUNCTION.
 RX MEDLINE=9225128; PubMed=1373391;
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;
 RT "An interferon-induced protein with release factor activity is a
 RT tryptophanyl-tRNA synthetase.";
 RL FEBS Lett. 300:162-166(1992).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- INDUCTION: BY INTERFERON GAMMA.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M77804; AAA67324.1; -
 DR EMBL: X59892; CAA42545.1; -
 DR EMBL: M61715; AAA61298.1; -
 DR EMBL: X62570; CAA44450.1; -
 DR EMBL: S82905; AAB39381.1; -
 DR EMBL: X67920; CAB94198.1; -
 DR EMBL: X67921; CAB94198.1; JOINED.
 DR EMBL: X67922; CAB94198.1; JOINED.
 DR EMBL: X67923; CAB94199.1; -
 DR EMBL: X67924; CAB94199.1; JOINED.
 DR EMBL: X67925; CAB94199.1; JOINED.
 DR EMBL: X67926; CAB94199.1; JOINED.
 DR EMBL: X67927; CAB94199.1; JOINED.
 DR EMBL: X67928; CAB94199.1; JOINED.
 DR PIR: A41706; A41706.
 DR PIR: A41633; A41633.
 DR PIR: JH0533; JH0533.
 DR PIR: S19246; S19246.
 DR Aarhus/Chent-2DPAGE; 3524; IEF.
 DR PDBT-2DPAGE; P23381; -
 DR MIM: 191050; -
 DR InterPro: IPR00738; WHEP-TRS.
 DR InterPro: IPR003035; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1tp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR PRINTS: PRO1039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 DR Antisacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 19 64 WHEP-TRS.
 FT SITE 164 173 "HIGH" REGION.
 FT SITE 349 353 "KMSKS" REGION.
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).
 FT CONFLICT 424 424 A -> R (IN REF. 4).
 SQ SEQUENCE 471 AA; 53165 MW; E9634444905A0D0 CRC64;

Query Match 99.3%; Score 2231; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 8e-178;
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MSYKAAGEDYKADCPGCPNPAPTSNGEPDTEAEEDVDYPTVOTSSAKGIDYOKLYRF 60
 DB 48 MSYKAAGEDYKADCPGCPNPAPTSNGEPDTEAEEDVDYPTVOTSSAKGIDYOKLYRF 107
 QY 61 GSSKIDELIRIRATGQRPHHLLRGIFPSHRDMOVLDAYENKKPFYLTGRGSSSE 120
 DB 108 GSSKIDELIRIRATGQRPHHLLRGIFPSHRDMOVLDAYENKKPFYLTGRGSSSE 167
 QY 121 AMHGHILPFIETKMLQDVFNPLVIGMTDDEKYLKMDLTLDQAYGDAVENAKDIIICGF 180
 DB 168 AMHGHILPFIETKMLQDVFNPLVIGMTDDEKYLKMDLTLDQAYGDAVENAKDIIICGF 227
 QY 181 DINKTFISDLDYMGSSGFKNVYKIOKHYTPNOVKCIFSTSDICGKISPAIQAP 240
 DB 228 DINKTFISDLDYMGSSGFKNVYKIOKHYTPNOVKCIFSTSDICGKISPAIQAP 287
 QY 241 SESNSFQIFRDRIIDICLIPCAIDQDYPFRMTDVAAPRIGPKPALHSTFFPALGQAO 300
 DB 288 SESNSFQIFRDRIIDICLIPCAIDQDYPFRMTDVAAPRIGPKPALHSTFFPALGQAO 347
 QY 301 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGRDTIEHRQFGNCDDVVSFMYLTFFFL 360
 DB 348 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGRDTIEHRQFGNCDDVVSFMYLTFFFL 407
 QY 361 EDDDKLEQIRKDYSGAMLTGELKALIEVLOPLAEHQARKKEYTDEIYKEFMTPKRLS 420
 DB 408 EDDDKLEQIRKDYSGAMLTGELKALIEVLOPLAEHQARKKEYTDEIYKEFMTPKRLS 467
 QY 421 FDFQ 424
 DB 468 FDFQ 471
 RESULT 2
 SYW_BOVIN STANDARD; PRT; 475 AA.
 ID SYW_BOVIN
 AC P17248;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (Ec 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN WARS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Retina;
 RX MEDLINE=91329348; PubMed=1907847;
 RA Garet M., Pajot B., Trezeguet V., Labouesse J., Merle M.,
 RA Gandar J.-C., Benedetto J.-P., Sallafrance M.-L., Alterio J.,
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to
 RT prokaryotic synthetases but near identity with mammalian peptide
 RT chain release factor.";
 RL Biochemistry 30:7809-7817(1991).
 RP [2]
 RP SEQUENCE OF 17-475 FROM N.A.
 RC TISSUE=Pancreas;
 RA Garet M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., la Bouessec B.,
 RA Labouesse J., Bonnet J.;
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 7.77387 Seconds

(without alignments)
2111.829 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEMTPKLSFDQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1 SYW_HUMAN	P47372 mycoplasma
2	2134.5	95.0	475	1 SYW_BOVIN	P46579 caenorhabditis
3	2056	91.5	475	1 SYW_RABIT	P07166 clostridium
4	2032	90.5	481	1 SYW_MOUSE	P21656 bacillus su
5	1210	53.9	395	1 SYW_SCHRO	P57602 buchnera ap
6	1163	51.8	432	1 SYW_YEAST	O57834 methanococ
7	803	35.8	385	1 SYW_PYRAB	O53836 mycobacteri
8	409.5	18.2	370	1 SYW_METJA	O92d76 rickettsia
9	397.5	17.7	364	1 SYW_METJA	P00954 escherichia
10	386	17.2	134	1 SYW_ENCCU	P57956 pasteurella
11	370.5	16.5	420	1 SYW_ARCFU	O49900 mycobacteri
12	353	15.7	374	1 SYW_AERPE	P73655 synechocyst
13	192	8.5	323	1 SYW_ARCFU	
14	175.5	7.8	341	1 SYW_CLOLO	
15	172	7.7	394	1 SYW_YEAST	
16	170.5	7.6	366	1 SYW_SULSO	
17	169	7.5	395	1 SYW_AQUAE	
18	166.5	7.4	346	1 SYW_CHLTR	
19	163.5	7.3	353	1 SYW_BORBU	
20	159.5	7.1	337	1 SYW_TREPA	
21	158	7.0	346	1 SYW_CHLMU	
22	152	6.8	344	1 SYW_CHLPN	
23	146	6.5	326	1 SYW_HELPY	
24	138	6.1	319	1 SYW_METJA	
25	136.5	6.1	528	1 SYW_BOVIN	
26	135.5	6.0	328	1 SYW_THEMA	
27	133	5.9	328	1 SYW_BACST	
28	132.5	5.9	343	1 SYW_MYCLE	
29	132.5	5.9	528	1 SYW_HUMAN	
30	130	5.8	326	1 SYW_HELPJ	
31	129	5.7	401	1 SYW_SCHPO	
32	128.5	5.7	379	1 SYW_YEAST	
33	128	5.7	334	1 SYW_HAEIN	

ALIGNMENTS

RESULT 1	ID	SYW_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381		P78535: Q9UDL3;		
DT	01-NOV-1991		(Rel. 20, Created)		
DT	01-DEC-1992		(Rel. 24, Last sequence update)		
DT	16-OCT-2001		(Rel. 40, Last annotation update)		
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)				
DE	(TrpRS) (IIP53) (hmRS).				
CN	WARS.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92105071; Pubmed=1761529;				
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;				
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human				
RT	fibroblasts.";				
RL	J. Biol. Chem. 266:24245-24248(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92107982; Pubmed=1763065;				
RA	Flecker J., Rasmussen H.H., Justesen J.;				
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa				
RT	protein (gamma 2) highly homologous to rabbit peptide chain release				
RT	factor and bovine tryptophanyl-tRNA synthetase.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92112058; Pubmed=1765274;				
RA	Prolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovjeva O.L.,				
RT	Kisselev L.L.;				
RT	"Cloning and nucleotide sequence of the structural gene encoding for				
RT	human tryptophanyl-tRNA synthetase.";				
RL	Gene 109:291-296(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92164636; Pubmed=1537332;				
RA	Buwaitt U., Flohit T., Boettger E.C.;				
RT	"Molecular cloning and characterization of an interferon induced				
RT	human cDNA with sequence homology to a mammalian peptide chain				
RT	release factor.";				
RL	EMBO J. 11:489-496(1992).				
RN	[5]				
RP	SEQUENCE OF 1-13 FROM N.A.				
RX	MEDLINE=9631994; Pubmed=8724762;				
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;				
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-				
RT	tRNA synthetase gene.";				
RL	MOL. BIOL. (Mosk) 30:319-329(1996).				
RN	[6]				
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.				
RC	TISSUE-Sperm;				
RX	MEDLINE=93292992; Pubmed=7685728;				

DB 342 IAMEKLERFLAEHOSRLEKAKTIANKLVEPPR 373

Search completed: October 24, 2002, 12:53:53
Job time : 16.2521 secs

Db 331 ECKNNTAEFRKPFELSVKREK 353

RESULT 13

T43806

tryptophan--tRNA ligase (EC 6.1.1.2) [Imported] - Encephalitozoon cuniculi (fragment)

C:Species: Encephalitozoon cuniculi

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43806

R:Peptatallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.

Mol. Biol. Evol. 15, 683-689, 1998

A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gen

A:Reference number: 226933; MUID:98277683

A:Accession: T43806

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PEY>

A:Cross-references: EMBL:AF012470; PIDD:CAA10034.1

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

C:Keywords: ligase

Query Match 17.2%; Score 386; DB 2; Length 134;

Best Local Similarity 51.3%; Pred. No. 3,Je-24;

Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

Qy 35 EDFVDPWTVOTSSAK----GIDYDKLIVRFSSKIDKELIRIRATGQRPNHFLRGIF 90

Db 3 EQRTTPDVEVSTDEVPALVDYKIINQFCGFNOALDRKLISKRAHYFRRGIV 62

Qy 91 FSHRDMQVLDAYENKKPFYLYTGRGSSSEAMHGHLPFTKWLQDVNPVLVQMTD 150

Db 63 FAHDFPMLLDELAINNPFLYLYTGRGSSKTMHIGHTIPFLCKYQMDAFKIRLITQTD 122

Qy 151 DEKLMKDLTID 162

Db 123 DEKLMKSMRLE 134

RESULT 14

E69461

tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C:Accession: E69461

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kinkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: utterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343

A:Accession: E69461

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 <KLE>

A:Cross-references: GB:AE00096; GB:AE000782; NID:92689309; PIDD:AA8954.1; PID:9264885

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 16.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 2,9e-22;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

Qy 38 VDPPTVOTSSAKGIDYDKLIVRFSSKIDKELIRIRATGQRPNHFLRGIFFSHRDM 97

Db 3 VTPPEVEGV---IDYSKLIEFGMOPF-SEVLPEID---NPHILMRGALIFGHRDYW 52

Qy 98 QVLDAYENKKPFYLYTGRGSSSEAMHGHLPFTKWLQDVNPVLVQMTDDEKYLK 157

Db 53 RIIFAMKKREKWMVMSGFMSG-LPHFGHKMTMDEIYMHOSAGKAV-AIADDEAHSVR 110

Qy 158 DLTLDQAYGAVENAKDIIACGFDINKTFFISDLDMGMSGFYKNV-KIOKHVTENOV 216

Db 111 GLSWKTRTELGMLYIKSLIALGLREDAVIYPOS-----KSHVKDLAEFLSAEVNSEL 164

Qy 217 KGIFGPTDSDICGISPPAIOAARFSSFPQIRDRDIOCLIPCAIDDPYRMRDY 276

Db 165 RAIFGMSDTSIARMEVTAIOADIL---HPQLSDFGPKPVVAVGADDPHMRITRDL 221

Qy 277 APR----- 280

Db 222 AARISIFEPVEGVAVRSKGAEYLSLRLDFEDDKIYEHNHDIGEAEIFERAVRKI 281

Qy 281 -----GYKPALHSTFPFALOGAQTMSASDPNSSIFLTDAKOIKTKVKNHAFSG 333

Db 282 EVEIGCAFIPSSSTYHRTTGLTG- -KMSSSKPEYSILIDPDEGAKVKM-AFTGC 338

Qy 334 RDTIEHRQPGCNDVDVSPFYLTFPLED-DDKLEQIRKQVTSQAMLTGELKALLIEVLQ 392

Db 339 RATAEQRRLGGEPPDCVVELYSFHLIDSELNQIEACREGRLCGCKMAAEELVK 398

Qy 393 PLIAHQARREK 405

Db 399 SLKEHOKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

Am. Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: D72477

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KAW>

A:Cross-references: DDBJ:AP000064; NID:95105945; PIDD:BA81476.1; PID:95106165

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2461

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 15.7%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 6,5e-21;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

Qy 38 VDPPTVOTSSAKGIDYDKLIVRFSSKIDKELIRIRATGQRPNHFLRGIFFSHRDM 97

Db 8 LDPW---GAVEIKDYDLRLTFTGIRPSEVL-PLRKAGMEPSFLMRGIIFGHRDFO 61

Qy 98 QVLDAYENKKPFYLYTGRGSSSEAMHGHLPFTKWLQ-DVENPVLVQMTDDEYLLW 156

Db 62 KILEAKARGEVAALVTFMPSGK-FHFGHKLTVDQLIYLQNGKRV--FVAIADAFAFV 118

Qy 157 KDLTLDQAYGAVEN-AKDIACGFDINKT-FISDLDMGMSGFYKNV-KIOKHVTEN 214

Db 119 RIRIGREAVRIAEVEYIANNMIALGLDPRKDEEYFQ---TNKGIPTVRLIOLSGVYAA 174

Qy 215 QVKGIFG-FPDSDCIGKISPPAIOAARFSSFPQIRDRDIOCLIPCAIDDPYRMR 273

Db 175 EMEAIYELIPRAKMAASLT---QAADILHVGQDEYGYR--HVVVPVAGADDPHRLT 227

Qy 274 RDAVAPR---IGYKPALHSTFPFALOGAQTMSASDPNSSIFLTDAKOIKTKVKNH 329

Db 228 RDLADRMAGVLELPRPASTYHKLQPGLDG--RKSSSRPSTFTLTPRPEVARNKLER-A 284

Qy 330 FSGGRDITIEHRQPGCNDV-DVSFMYLTFLEDDDKLEQIRKQVTS---GAMLTGELK 385

Db 285 LTGGRATAEQRRLGGEVPCSVYHMDLYHLMRPDGEVKNH---YTSIRLGLITIGCECKO 341

Qy 386 ALIEVLQPLIAEHQARREKVTDEIVKEMTPR 417

[illegible][illegible]

```

Qy 215 QVKFEGTSDSCJGKISFPAIOAAPSNSNFQIDFRDRT--DIQCLIPCAIDDDPYFRM 272
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 168 EMKALYGGKGTGTHGHVAPAPIVQAVDIL--HPQIDENLSEPKPVVYPVQIDPHRL 224
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 273 TRDVAPR--IGYKPKALLHSTFFPALOGAOTKKASADPNSSIFLTDTAKOIKTVKNHA 329
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 225 TRDIANRAKKEKFFIPSSYTHRFYMGLLDG--KSSSSPEPAIFLTDEKVKKKIIS-A 281
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 330 FSGGHDTEIEHRQFCG--NCDVDVSNWYLTFFLEDDKLEQIKRDYTSGAMLTGELKA 386
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 282 KTGSGRETELEHKKGGVGVVEECVVELFLY--HLILDKEIAEIVOKRSGSGELTGCKCKM 339
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 387 LIEVQPLIAEHQARKREVTDEIYK 411
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 340 AYERVVEFLDKLKEKREQAKETIAVK 364
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

QY	382	ELKALIEVLOPLIAHQARRK 404
QY	17.7%	Score 397.5; DB 2; Length 364;
QY	27.9%	Pred. No. 1.5e-24;
QY	107;	Conservative 76; Mismatches 153; Indels 47; Gaps 14;
QY	2	IDPW-----GSAK-LEXPDLIEFGVRR-SEVLDEV-----PESSWLMRRRIILIGHNRYE 50
QY	98	OVLDAVENKKRPYYLTGSGPSSSEAMNHVGLIPLEFTKKLQDVFNPLVYIOMTDDKYLWK 157
QY	51	RISAMKKGGEFPAVYTGMPSPGR-NHIGKKMIVDQLRN-YRMCAGELPIPIADMEAVSAR 108
QY	158	DITLQAGDAVEN-AKQIICAGPDK-----TFESDLDDYGMSSGGYKKNVYK 206
QY	109	GDFEDSRRIALIEEYIAGYIALGDLEKQNIHVYLQSNLWVEDLAYV----- 156
QY	207	IOKHVTFNQVAGIGFETSDSGIKTISFPALQAAVSFNSPFOIFRDRYTDIOLPCALDQ 266
QY	157	LAGKVNFEHLRIAYGFTGSTSAMHYAPRIQVSDILHPQDLDELGPR--PVIVPVGDQ 213
QY	267	DPIYFMTQDVAPRI-----GYKPPALHSTFFPPALQAGOTKKASADPNNSITLPTAKQIK 322
QY	214	DPIHLTFLDIARAFDRYGFLLPSSTYHRHFGGLTG--KMSNPKSAIFLSDTPEAE 271
QY	323	TKVNKHAHSGSGDTEEHROFGNCDVVSFMYLTFLE--DDDKLEQIRKYDTSGAMLTG 381
QY	272	AKI-RNATTCGHEITLKEORELGGVPEECITITETLLYHMSGSDSRLEETIYESCRRGTLMCG 330

A:Residues: 1-212, 'GD', 215-471 <FRO2>
 A:Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368
 A:Experimental source: fibroblast
 C:Genetics:
 A:Gene: GDB:WARS; IFF53
 A:Cross-references: GDB:119632; OMIM:191050
 A:Map position: 14q23-14q31
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 1.4e-173;
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MSYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYDKLIVRG 60
    |||||
DB 48 MSYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYDKLIVRG 107
    |||||

OY 61 GSKIDKELINRIERATGQRPHRLRGIFFSHRDMNOVLDAEYENKPPYLTYGRGPSSA 120
    |||||
DB 108 GSKIDKELINRIERATGQRPHRLRGIFFSHRDMNOVLDAEYENKPPYLTYGRGPSSA 167
    |||||

OY 121 AMHVGHLIPFTKWLQDVFNPLVIOMTDDEKYLKDLTLDQAYGDAVENAKDIACGF 180
    |||||
DB 168 AMHVGHLIPFTKWLQDVFNPLVIOMTDDEKYLKDLTLDQAYGDAVENAKDIACGF 227
    |||||

OY 181 DINKTFISDLDYMGSSGFEYKNVYKIQKHVFNQVKGIFGFTSDSDCIGKISFPALQAA 240
    |||||
DB 228 DINKTFISDLDYMGSSGFEYKNVYKIQKHVFNQVKGIFGFTSDSDCIGKISFPALQAA 287
    |||||

OY 241 SFSNFPQIFRDRTOICLPCAIIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAG 300
    |||||
DB 268 SFSNFPQIFRDRTOICLPCAIIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAG 347
    |||||

OY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFAFGSGRDTIEEHRQFGNCVDVSMYLTFFL 360
    |||||
DB 348 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFAFGSGRDTIEEHRQFGNCVDVSMYLTFFL 407
    |||||

OY 361 EDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDVYKEMPRKLS 420
    |||||
DB 408 EDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDVYKEMPRKLS 467
    |||||

OY 421 FDFEQ 424
    |||||
DB 468 FDFEQ 471
    |||||

```

RESULT 2

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - bovine

N:Alternate names: tryptophanyl-tRNA synthetase

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 21-Jul-2000

C:Accession: A40279; JN0354; S10460; S14540

R:Garret, M.; Pajot, B.; Tirezneet, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti

Biochemistry 30, 7809-7817, 1991

A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic s

A:Reference number: A40279; MUID:91329348

A:Accession: A40279

A:Residues: 1-475 <GAR>

A:Molecule type: mRNA

A:Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799

A:Experimental source: pancreas

A>Note: The authors translated the codon CTG for residue 347 as Ala and CAG for residue

R:Zarganova, T.A.; Kovaleva, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.

Bioorg. Khim. 15, 1307-1311, 1989

A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca

A:Reference number: JN0354; MUID:90211408

A:Accession: JN0354

A:Molecule type: protein

A:Residues: 112-124;282-287, 'N',288, 'F',289-292, 'Q',293-294, 'IR',336-353;423-441,443-
 A:Experimental source: liver
 A>Note: this paper is in Russian
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;
 Best Local Similarity 95.0%; Pred. No. 1e-165;
 Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

```

OY 2 SYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYDKLIVRG 61
    |||||
DB 54 SYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVOTSSAKGIDYDKLIVRG 113
    |||||

OY 62 SSKIDKELINRIERATGQRPHRLRGIFFSHRDMNOVLDAEYENKPPYLTYGRGPSSA 121
    |||||
DB 114 SSKIDKELINRIERATGQRPHRLRGIFFSHRDMNOVLDAEYENKPPYLTYGRGPSSA 173
    |||||

OY 122 MHVGHILIPFTKWLQDVFNPLVIOMTDDEKYLKDLTLDQAYGDAVENAKDIACGF 181
    |||||
DB 174 MHVGHILIPFTKWLQDVFNPLVIOMTDDEKYLKDLTLDQAYGDAVENAKDIACGF 232
    |||||

OY 182 INKTFISDLDYMGSSGFEYKNVYKIQKHVFNQVKGIFGFTSDSDCIGKISFPALQAA 241
    |||||
DB 223 INKTFISDLDYMGSSGFEYKNVYKIQKHVFNQVKGIFGFTSDSDCIGKISFPALQAA 292
    |||||

OY 242 FSNSEFPQIFRDRTOICLPCAIIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAG 301
    |||||
DB 293 FSNSEFPQIFRDRTOICLPCAIIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAG 352
    |||||

OY 302 KMSASDPNSSIFLDTAKQIKTKVKNKHAFAFGSGRDTIEEHRQFGNCVDVSMYLTFFL 361
    |||||
DB 353 KMSASDPNSSIFLDTAKQIKTKVKNKHAFAFGSGRDTIEEHRQFGNCVDVSMYLTFFL 412
    |||||

OY 362 DDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDVYKEMPRKLS 421
    |||||
DB 413 DDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDVYKEMPRKLS 472
    |||||

OY 422 DFO 424
    |||||
DB 473 DFO 475
    |||||

```

RESULT 3

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - rabbit

N:Alternate names: tryptophanyl-tRNA synthetase

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Sep-1992 #sequence_revision 13-Feb-1998 #text_change 26-May-2000

C:Accession: A35904; S37396

R:Lee, C.C.; Craigen, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.

Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990

A:Title: Cloning and expression of a mammalian peptide chain release factor with sequ

A:Reference number: A35904; MUID:90239043

A:Accession: A35904

A:Molecule type: mRNA

A:Residues: 1-475 <LEE>

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

A:Molecule type: mRNA

A:Cross-references: GB:M33460

A:Residues: 166-177 <FRO>

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 14.2521 Seconds
(without alignments)
2858.658 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471
Perfect score: 2246
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEFMTPRKLSFDQ 424

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1 A41706	tryptophan--trna 1
2	2134.5	95.0	475	1 YWBO	tryptophan--trna 1
3	2041	90.9	475	1 TWBPR	tryptophan--trna 1
4	2032	90.5	481	2 S50053	tryptophan--trna 1
5	1210	53.9	395	2 S58157	hypothetical prote
6	1163	51.8	432	2 S51901	tryptophan--trna 1
7	907	40.4	385	2 C90190	tryptophan--trna 1
8	803	35.8	385	2 C75020	tryptophan--trna 1
9	626.5	27.9	301	2 G71206	tryptophan--trna 1
10	534.5	23.8	380	2 G84373	tryptophan--trna 1
11	409.5	18.2	370	2 E64476	tryptophan--trna 1
12	397.5	17.7	364	2 E69131	tryptophan--trna 1
13	386	17.2	134	2 T43806	tryptophan--trna 1
14	370.5	16.5	420	2 E69451	tryptophan--trna 1
15	353	15.7	374	2 D72477	probable tryptoph
16	274.5	12.2	513	2 F84371	tryptophan--trna 1
17	192	8.5	323	2 H69346	tyrosyl--trna synth
18	188	8.4	364	2 D72512	probable tyrosyl-t
19	185.5	8.3	341	2 D95260	tryptophan--trna 1
20	185.5	8.3	341	2 G98125	tryptophan--trna 1
21	174.5	7.8	341	2 B86633	tryptophan--trna 1
22	172	7.7	394	2 A45999	tyrosine--trna 1
23	170.5	7.6	366	2 S75410	tyrosine--trna 1
24	169	7.5	395	2 H70385	tryptophan--trna 1
25	167	7.4	351	2 E75438	tryptophan--trna 1
26	166.5	7.4	346	2 B71496	tryptophan--trna 1
27	164.5	7.3	895	2 A86410	protein F3M18.22
28	163.5	7.3	353	2 E70100	tryptophan--trna 1
29	163	7.3	460	2 C84750	probable tyrosyl-t

30	159.5	7.1	337	2 F71300	tryptophan--trna 1
31	158.5	7.1	375	2 B75072	tyrosyl--trna synth
32	158	7.0	346	2 C41654	tryptophan--trna 1
33	154.5	6.9	102	2 T74994	probable tryptoph
34	152	6.8	344	2 H86590	tryptophan--trna 1
35	152	6.8	344	2 C72034	tryptophan--trna 1
36	149.5	6.7	327	2 C84374	tyrosyl--trna synth
37	149.5	6.7	408	2 T03741	probable tyrosine-
38	146	6.5	339	2 E64676	tryptophan--trna 1
39	146	6.5	375	2 F71093	tyrosine--trna 1
40	145	6.5	337	2 A71066	tryptophan--trna 1
41	138	6.1	319	2 H69102	tyrosine--trna 1
42	135.5	6.0	328	2 C72370	tryptophan--trna 1
43	135	6.0	365	2 E82052	tryptophan--trna 1
44	132.5	5.9	343	2 S73024	tryptophan--trna 1
45	132	5.9	328	1 YWBSF	tryptophan--trna 1

ALIGNMENTS

RESULT 1

A41706
tryptophan--trna 1 ligase (EC 6.1.1.2) [similarity] - human
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor h
C:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text change 19-May-2000
C:Accession: A41633; A41706; S19246; JN0676; JH0533; S26287
R: Fleckner, J.; Rasmussen, H.H.; Justesen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9
A:Reference number: A41633; MUID:92107982
A:Accession: A41633
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821
R: Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 24245-24248, 1991
A:Title: Interferon induces tryptophan--trna synthetase expression in human fibrobla
A:Reference number: A41706; MUID:92105071
A:Accession: A41706
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:M77804; NID:9184656; PIDN:AAA67324.1; PID:9184657
R: Buwilt, U.; Fliohr, T.; Boeltger, E.C.
EMBO J. 11, 489-496, 1992
A:Title: Molecular cloning and characterization of an interferon induced human cDNA w
A:Reference number: S19246; MUID:92164636
A:Accession: S19246
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-423, 'R', 425-471 <BUW>
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709
A:Note: 213-Ser and 214-Tyr were also found
R: Frolova, L.Y.; Grigorjeva, A.Y.; Sudomolina, M.A.; Kisselev, L.L.
Gene 120, 237-245, 1993
A:Title: The human gene encoding tryptophan--trna synthetase: Interferon-response el
A:Reference number: JN0676; MUID:93292992
A:Accession: JN0676
A:Molecule type: DNA
A:Residues: 1-141,182-471 <FRO>
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X6792
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB
A:Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residu
R: Frolova, L.Y.; Sudomolina, M.A.; Grigorjeva, A.Y.; Zhirovleva, O.L.; Kisselev, L.L.
Gene 109, 291-296, 1991
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr
A:Reference number: JH0533; MUID:92112058
A:Accession: JH0533
A:Molecule type: mRNA

Fri Oct 25 09:23:05 2002

us-09-813-718-10_copy_48_471.ra

Page 9

Search completed: October 24, 2002, 12:54:33
Job time : 11.3652 secs

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-100-4

Query Match
Best Local Similarity 25.3%; Pred. No. 0.00081;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG--NAKMSK 61

QY 306 SDPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSFWLTFE--LEDD 363
DB 62 S-LNNGIYLADDDTLRKTKVMSWTDPDHLRVEDPGKIEGN---WVFHYLDVGRPEDA 116

QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQAARREYVDEI 409
DB 117 QEADMKERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162

RESULT 12
US-09-183-134-2
Sequence 2, Application US/09183134
Patent No. 6165759
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,134
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/923,867
FILING DATE: 04-SEP-1997
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-183-134-2

Query Match
Best Local Similarity 25.3%; Pred. No. 0.00081;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG--NAKMSK 61

DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG--NAKMSK 61

QY 306 SDPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSFWLTFE--LEDD 363
DB 62 S-LNNGIYLADDDTLRKTKVMSWTDPDHLRVEDPGKIEGN---WVFHYLDVGRPEDA 116

QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQAARREYVDEI 409
DB 117 QEADMKERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162

RESULT 13
US-09-492-581-4
Sequence 4, Application US/09492581
Patent No. 6346409
GENERAL INFORMATION:
APPLICANT: Gentry, Danlie
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6346409el trps
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,581
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-492-581-4

Query Match
Best Local Similarity 25.3%; Pred. No. 0.00081;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG--NAKMSK 61

QY 306 SDPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSFWLTFE--LEDD 363
DB 62 S-LNNGIYLADDDTLRKTKVMSWTDPDHLRVEDPGKIEGN---WVFHYLDVGRPEDA 116

QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQAARREYVDEI 409
DB 117 QEADMKERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162

EARLIER APPLICATION NUMBER: 60/092,866
 EARLIER FILING DATE: JULY 15, 1998
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 28
 LENGTH: 377
 TYPE: PMT
 ORGANISM: *Synechocystis* sp.
 US-09-352-990-28

Query Match 4.9%; Score 109; DB 4; Length 377;
 Best Local Similarity 21.0%; Pred. No. 0.0018;
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;

QY 91 FSHRNMNVLAAYENKKPFYLYTGRGSPSEAMHVGHLIPFTKYLQOVENVPLVQMTD 150
 DB 33 FAHSTTM-----DKPRIL-SGVQPTGN-LHLGNVIGAI-RSWEQ-----QQHY 75
 QY 151 DEKYLAKDL-----TLDAQYDAVENAKDIIACGFDINKTFIF----- 188
 DB 76 DNEFCVVDLHATVPHPNQTIAQ---DPLTIALYLAGIDQYTFIVQSHVAHSELA 132
 QY 189 -----SGLDWGSSGGYKKNVYKQKHYTVNQVKGIFGTFDSCIGKISPPAIDAAASF 242
 DB 133 WLNCVTEPLNLMERMIQEKERAKVKGGENVS-----VGLLDYPLVMAA--- 174
 QY 243 SNSPQIFRDRDIQCL---IPCAIDDDPYFRMTRDY-----APRIGYPKPA 286
 DB 175 -----DILLYDADKYPVEDOKOHELTRDIYIRINDKFGEDAFVLPEPL 222
 QY 287 L-LHSTFPALOGAQTKNASADPN--SSIFLYDTAKQIKTKVKNHAFSGGRDTEHROF 343
 DB 223 IRKEGARVMSLADGTKKMSKSESELRLNLDPEMIKKVKK----- 266
 QY 344 GCNCVVDVSEMYLTFLEDDDKLEIRKDYTSGAMLTEGELKAL----- 387
 DB 267 ---CKTDPO-RGLWF---DDPERPECHNLLTYLTLNMQTEKAVAOEACMGWGQFKPL 319
 QY 388 -----IEVLOPLIAEQARKREVTDEIVK 411
 DB 320 TETATALEPIQAKYAEILADRGEL-DRILQ 349

RESULT 10
 US-08-923-867-2
 Sequence 2, Application US/08923867
 Patent No. 5851809

GENERAL INFORMATION:
 APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/923,867
 FILING DATE: 04-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9619072.3
 FILING DATE: 12-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimm, Edward R
 REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P31624
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-923-867-2

Query Match 4.8%; Score 108; DB 2; Length 197;
 Best Local Similarity 25.3%; Pred. No. 0.00081;
 Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPPALLHSTFPALOGAQTKMSA 305
 DB 5 VPVGTDQKPMIEQPREIYRSFNNAVNCVLYVEPEIYENE--RAGRLPGLDG-NAKMSK 61
 QY 306 SDPNSITPLTAKQIKTKVKNHAFSGGRDTIEHROFGWCVDVSPMYLTF--LEDD 363
 DB 62 S-LNNGIYLAADADDTLRKKVSMYTPDPHIVEDEPGKIEGN---MFHYLYDVGREPDA 116
 QY 364 DKLEIRKDYTSGAMLTEGELKALIEVLOPLIAEQARKREVTDEI 409
 DB 117 OEIADMKERYORGIGDYKTKRYLLELIERLGPTRERIEFADKM 162

RESULT 11
 US-08-928-100-4
 Sequence 4, Application US/08928100
 Patent No. 6046174

GENERAL INFORMATION:
 APPLICANT: Gentry, Danile
 APPLICANT: Greenwood, Claire
 APPLICANT: Lawlor, Elizabeth
 TITLE OF INVENTION: NO. 6046174el trps
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,100
 FILING DATE: 12-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9619072.3
 FILING DATE: 12-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimm, Edward R
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P31624-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/705,868
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0117 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 388 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 1184699
 US-09-123-615-4

Query Match 5.98; Score 132.5; DB 3; Length 388;
 Best Local Similarity 21.28; Pred. No. 4.7e-06;
 Matches 79; Conservative 64; Mismatches 149; Indels 80; Gaps 19;

QY 83 HFLRGII--FESHDMNVLDAYENKKRFVLYTGRGSPSEAMHGHILPFI--FTKWLQDV 139
 DB 12 HLIRNLOEVLGEERKILKERELK---IYWGATATGKP-HAAYVPMKSIADFLKAG 66
 QY 140 FNVPLVI---QMTDDEKYLKDLTDQAYGDAVENAKDIACGPDINKTFISDLDMG 195
 DB 67 CEVILFADLAHLYDNMKAPELLELRYSYENVIKAM-LESIGVPLEKLFKIGTDQ- 124
 QY 196 MSSGFYKAVKVIQKHVFNQVK---GTFGTTSDSCIGKISFPALQAPSPNSPEQIFR 251
 DB 125 LSKETTLDVYRLSSVYVQHDSSKKAQAEVYKQVEHPLLSGLLYPGIOA---LDEEYLKV- 179
 QY 252 DRTIOGICLIPCAIDODPFERMTROVAPRIGYKPKALHSTFFPALQAGQTKMSASDPSS 311
 DB 180 ---DAQF---GGIDQRKFTFAEKLTPALGTSKRVHLMNPMVPLTG--SKMSSSEESK 231
 QY 312 IFLDTAKOIKTKYKNKNAF-----SGGRDTIEEH-----ROFGMCDVD 350
 DB 232 IDLDRKDKVKKLKK-AFCERGVNENNGVLSFKHVLFPKSEFVILDRKMGGN---- 286
 QY 351 VSEMYLFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLAEHQARKREYTDIY 410
 DB 287 ---KTYTAYVD-----LEKDFAAEVVHPGDLKNSVEVALNKLK-----DPFR 325
 QY 411 KEFMTF--RKLS 420
 DB 326 EKFMTPALKLKA 337

RESULT 8
 US-08-855-910-11
 Sequence 11, Application US/08855910
 Patent No. 6221640
 GENERAL INFORMATION:
 APPLICANT: Tao, Jianshi
 APPLICANT: Sassanfar, Mandana
 APPLICANT: Gallant, Paul L.
 APPLICANT: Shen, Xiaoyu
 APPLICANT: Avriuch, Anthony S.
 APPLICANT: Yu, Russell V.
 APPLICANT: Nair, Shamila
 TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
 NUMBER OF SEQUENCES: 71
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Militia Drive

CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/855,910
 FILING DATE: 14-May-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Brook, David E.
 REGISTRATION NUMBER: 22,592
 REFERENCE/DOCKET NUMBER: CP195-08
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 861-6240
 TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 418 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-855-910-11

Query Match 5.18; Score 113.5; DB 4; Length 418;
 Best Local Similarity 20.58; Pred. No. 0.00069;
 Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;

QY 111 LYTGRGSEAMHGHILPFI--FTKWLQDVENVPLVI----- 146
 DB 33 LYCGVDPGDSMHGHILPFI--FTKWLQDVENVPLVI----- 146
 QY 147 -----QMTDDEKYLW-KDLTLDQAYGDAVEN-----AKDIIA 177
 DB 93 AVQHNVDLSNQMKKLRGKDAEYTMVNNYDLSLDFLRDVGKNNVTMLAKDIVA 152
 QY 178 CGFDINKTFISDLDMGMS--GFYKAVKIQKHVFNQVKG----- 218
 DB 153 S--RLESGISFTFTYQIQLSIDFY---TLHKKNIIQLQIGAGQWGNITAGLDLIRKE 207
 QY 219 -----IRGEFT-----DSDCIGKISFPALQAPSPNSP--QIFRRTDIOGLPCA 264
 DB 208 GPEAKVFGTLTPLMKADGTRFGKTAGALWLDPKRTISPFETFPWLNQD----- 258
 QY 265 DODEYFRMTROVAPRIGYKPKALHSTFFPALQAGQTKMSASDPSSIFLTDTAKOIKTK 324
 DB 259 -----RDV---IKYLK-----FFFLDKEE-----IDALAEKYEKE 286
 QY 325 VNKHAFSGGRDTIEEHROFGNCVDVSEMYLFFLEDDDKLEQIRKDYTSGAMLTGELK 384
 DB 287 PKR--EAQRRLAE-----VTRFVHDDALAEBAQK--ISEALFSGNIK 326
 QY 385 KALIEVLQ-----PLAEHQARKREYTDIYKEFMTPK 418
 DB 327 DLTEEIEQGLEHVPY-ETTKAKNIVDWLVTIEIEPSK 365

RESULT 9
 US-09-352-990-28
 Sequence 28, Application US/09352990
 Patent No. 6255090
 GENERAL INFORMATION:
 APPLICANT: Famodu, Layo O.
 APPLICANT: Orozco, Buddy
 APPLICANT: Rafalski, Antoni
 TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
 FILE REFERENCE: BB-1191
 CURRENT APPLICATION NUMBER: US/09/352,990
 CURRENT FILING DATE: 1999-07-14

```

; INFORMATION FOR SEQ ID NO: 2:
;
;   SEQUENCE CHARACTERISTICS:
;
;     LENGTH: 409 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;
;     TOPOLOGY: linear
;
;   MOLECULE TYPE: protein
;
US-08-743-130A-2

```

Query Match	7.2%	Score 162.5;	DB 2,	Length 409;
Best Local Similarity	20.7%	Pred. No. 2.3e+09;		
Matches 78;	Conservative 59;	Mismatches 132;	Indels 107;	Gaps 16;

```

OY 99 VLDVAENK-KFPLYUTRGPSSEMAHVGHILPF-----FTKWLDV-----F 140
Db 27 IKDVLKENRNVKLYWGAFTAPGR-HCGYVPMKTLAHLKACGEVYLLADLHAFLDNK 85
OY 141 NVPLVIOMTDEKYLMDLTLDDQYGDAVENAKDIACGFIDKNTFFESLDYMGSSGF 200
Db 86 KALEVYVYRAKYEEFVVKALTKLSINPIEBLKFPVSGSSYQKGADY---MDLFKLSINIV 142
OY 201 YKNVVK-----IQKHVTFNVOYKGFJFGTSDSDCIKISFPALIOAPRSTNSFPIFRDRD 255
Db 143 SQNDAKRAGADYVVOVANPLLSGLI-----YPLMOA-----IDEBHLG 180
OY 256 IQCLIPCAIDDDPFYRMTBDVAPRIPKRALHSTFPALOGAQTMSADPNSISPLT 315
Db 181 VDAOFG-GVDORAKFTVLAEEHLPSIGTKRAHLMNPVPL-GQGGKMSASDPNSKIDII 238
OY 316 DTAQOIKTKYVKNHAFSGS-RDT-----IEEHROFGN 346
Db 239 EEPYVKKKVNASACARGELKDNLIAFIEVYIOPIALKLTGVGAEKFLDIDPEKYG- 297
OY 347 CDVDFSEMYTLTFLEDDDKLEQIKDYTSGLMTGELKALI-----EYIOLPIAEHOARR 402
Db 298 ---PLSY-----DSIEQLKADVDGDKLAPPDKLGVAIDKINELLAPIRAEFESS- 343
OY 403 KEVYDEIKKEEMTPRK 418
Db 344 -----EEFOVAOK 351

```

RESULT 6
US-08-705-868-4
Sequence 4, Application US/08705868
Patent No. 5885798
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,868
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0117 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1184659

Query Match	5.9%	Score 132.5	DB 2	Length 388
Best Local Similarity	21.28%	Pred. No. 4.7e-06		
Matches 79, Conservative	64	Mismatches 149,	Indels 80,	Gaps 19

```

Qy      83 HLFIRGI--PFSHRDMONVDAYENKKRPVLYTGRGSPSSEAMHGHILPI--PTKMLDQV 139
Db      12 HILIRNLOEVLGEKKLKEILEREK-----IYGTATTCGR-PHYAVFVPSKLADELKAG 66
Qy      140 FNVPLVI-----QMTDDEKYLTKDITLTDQAYAGDAEYNAKDIAGCFDINTKFFIESDLDYMG 195
Db      67 CEVITLLEADLHAYLDNNKAPRHELLELKVSYEENIKRM--LESIGVPELEKLFIKGDYQ-- 124
Qy      196 MSSGFGYKNVAKIQKHVTFENQV-----GIFGFTSDCIGKISFPAIQAPFSFSNPQIFR 251
Db      125 LSKETVLLDVRILSSVYTOHDSKKGAEEVYKQVEHPRLISGLLYPGLOA--LDEEYLKV-- 179
Qy      252 DRTDQGLIPCAIDIDDPYFERTROVARIRGPKRALLHSFFPALQAGAOIKMSAOPNS 311
Db      180 ---DAQF---GDIQRIKIFFAEKYLLPALGSKRVHMLNPNVPGLTG--SKMSSSEEEK 231
Qy      312 IFLDTAKQIKTKYKNKHA-----SGGRDITREH-----ROEGGMDVD 350
Db      232 IDLLDRKEDKVDKKLKK-AFCERGVENNNGVLSFKHVLFLPKSRFVILRDEKNGN---- 286
Qy      351 VSFMYLFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVYDEIV 410
Db      287 ---KTYAYYVD-----LEKDFAAEYVHPDCLKNSYEVALNKL-----DPIR 325
Qy      411 KEFMTP--RKLS 420
Db      326 EKFNTPALKLA 337

```

```

: RESULT 7
: US-09-123-615-4
: Sequence 4, Application US/09123615
: Patent No. 6090377
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Coleman, Roger
: APPLICANT: Au-Young, Janice
: APPLICANT: Murry, Lynn E.
: TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: U.S.
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: Fastseq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/123, 615
: FILING DATE:
:

```


[illegible]

RESULT 2
US-08-928-100-2
; Sequence 2, Application US/08928100

```

GENERAL INFORMATION:
APPLICANT: Gentry, Danlle
APPLICANT: Greenwood, Claire
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6046174el trps
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-928-100-2

```

Query Match	8.3%;	Score 185.5;	DB 3;	Length 341;
Best Local Similarity	24.0%;	Pred. No. 4.7e+12;		

Matches	81;	Conservative	60;	Mismatches	128;	Indels	69;	Gaps	15;
QY	107	KEFYLYTGRCSSSEAHMHGHLPIFTYKWLQDENVNPLVYIQMTDDEKY-LMKDLTLDOAY	165						
Db	3	KEIL-TGDRIPGK-LIHGHVGSLEKNR-----VLLQEDSKYDMVEFLDQOAL	49						
QY	166	GVAVN-----AKDIIAGFDIDNKPIF--SDLDVGNSSSGYKVV---KQK	209						
Db	50	TDHADDPTIVESIGNVALDYAVGLDPKSTFIQSQIPETAEISMYMNLVSLARLER	109						
QY	210	HYTFNQVKGIEGFTSDSGIKSIFPAIOAA--PSFSNSFPQIFRDRTDIQCLIPCAIDQD	267						
Db	110	NPTVKTETLSQKFGESIPGLVLYIAQAADITARKAV-----VPGCTQOK	156						
QY	268	PYFRMRD-----VABRIGPKDALLHSFFPALGAGQTKMSADPNSSITF	313						
Db	157	PMIEOTREIVRSFNNAYNCDVLYEEGIIYENB--RAGRLPGLDQ-NAKMSKS-LNNGIY	212						
QY	314	LTDPTAKOTIKTKYKNAHAFSGRDTIEHHQFGNCQVDSFWYLTFF--LEDDDKLEQIRK	371						
Db	213	LADDDADTLTKKWKMSAKTDPDHLRVDDPGCITBEN---WVFHLVDFGRPEDAQETADMK	268						
QY	372	DYTSGAMLTGELKALIEVLOPLIIEHQARREYVDEI	409						
Db	269	KYQRGGLGDVKTTRKRYLLEIRELEGRIPINRKTIEFAKDM	306						

RESULT 3
US-09-492-581-2

```

1  GENERAL INFORMATION:
2  APPLICANT: Gentley, Danile
3  APPLICANT: Greenwood, Claire
4  APPLICANT: Lawlor, Elizabeth
5  TITLE OF INVENTION: No. 6346409e1 ttps
6  NUMBER OF SEQUENCES: 6
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: SmithKline Beecham Corporation
9  STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: PASTEDFO for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/492,581
21 FILING DATE:
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US/08/928,100
25 FILING DATE: 12-SEP-1997
26 APPLICATION NUMBER: 9619072.3
27 FILING DATE: 12-SEP-1996
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Gimm1, Edward R
30 REGISTRATION NUMBER: 38,891
31 REFERENCE/DOCKET NUMBER: P31624-1
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 610-270-4478
34 TELEFAX: 610-270-5090
35 TELLEX:
36 INFORMATION FOR SEQ ID NO: 2:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 341 amino acids
39 TYPE: amino acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 10.3652 Seconds
(without alignments)
999.159 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEFWTPRKLSPDFQ 424

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/laa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/laa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/laa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/laa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/laa/PCUS.COMB.pep.*
6: /cgn2_6/ptodata/1/laa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	54.3	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	8.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	8.3	341	US-09-492-581-2	Sequence 2, Appl
4	164.5	7.3	409	US-08-743-130A-39	Sequence 39, Appl
5	162.5	7.2	409	US-08-743-130A-2	Sequence 2, Appl
6	132.5	5.9	388	US-08-705-868-4	Sequence 4, Appl
7	132.5	5.9	388	US-09-123-615-4	Sequence 4, Appl
8	113.5	5.1	418	US-08-855-910-11	Sequence 11, Appl
9	109	4.9	377	US-09-352-990-28	Sequence 28, Appl
10	108	4.8	197	US-08-923-867-2	Sequence 2, Appl
11	108	4.8	197	US-08-928-100-4	Sequence 4, Appl
12	108	4.8	197	US-08-183-134-2	Sequence 2, Appl
13	108	4.8	197	US-09-492-581-4	Sequence 4, Appl
14	104.5	4.7	370	US-08-415-593-45	Sequence 45, Appl
15	95	4.2	418	US-08-844-054-2	Sequence 2, Appl
16	95	4.2	418	US-09-347-333-2	Sequence 8, Appl
17	94.5	4.2	877	US-08-907-166-8	Sequence 4, Appl
18	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
19	90.5	4.0	898	US-08-465-994C-4	Sequence 4, Appl
20	90.5	4.0	898	US-08-966-145-4	Sequence 4, Appl
21	90.5	4.0	920	US-08-101-593-4	Sequence 4, Appl
22	89.5	4.0	344	US-09-393-554-2	Sequence 2, Appl
23	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
24	89.5	4.0	898	US-08-465-994C-2	Sequence 2, Appl
25	89.5	4.0	898	US-08-966-145-2	Sequence 2, Appl
26	89.5	4.0	920	US-08-101-593-2	Sequence 2, Appl
27	89	4.0	719	US-08-082-849B-31	Sequence 31, Appl

28	89	4.0	719	5	PCT-US94-01624-31	Sequence 31, Appl
29	88	3.9	606	2	US-08-883-534-3	Sequence 3, Appl
30	88	3.9	606	2	US-09-204-764-3	Sequence 3, Appl
31	87.5	3.9	855	4	US-08-880-865A-10	Sequence 10, Appl
32	87	3.9	428	3	US-08-331-625A-43	Sequence 43, Appl
33	87	3.9	1452	3	US-08-331-625A-2	Sequence 2, Appl
34	87	3.9	1452	5	PCT-US93-04384-18	Sequence 18, Appl
35	87	3.9	1452	5	PCT-US93-04692-2	Sequence 2, Appl
36	86	3.8	1443	1	US-08-308-872B-2	Sequence 2, Appl
37	85	3.8	410	4	US-09-352-990-16	Sequence 16, Appl
38	83.5	3.7	510	1	US-08-249-112-3	Sequence 3, Appl
39	83.5	3.7	510	5	PCT-US95-06556-3	Sequence 3, Appl
40	83.5	3.7	970	1	US-08-375-709-7	Sequence 7, Appl
41	83.5	3.7	970	4	US-08-752-929-7	Sequence 7, Appl
42	83.5	3.7	970	4	US-09-090-793-5	Sequence 5, Appl
43	83	3.7	400	4	US-09-352-990-18	Sequence 18, Appl
44	83	3.7	551	1	US-08-484-493-15	Sequence 15, Appl
45	83	3.7	551	1	US-08-484-494-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-876-885-26
Sequence 26, Application US/08876885
Patent No. 6174713
GENERAL INFORMATION:
APPLICANT: Shen, Xiaoyu
APPLICANT: Houman, Fariba
TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPHANYL-TRNA
TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING
TITLE OF INVENTION: SAME
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,885
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI97-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-876-885-26

Query Match 54.3%; Score 1218.5; DB 4; Length 424;
Best Local Similarity 57.1%; Pred. No. 5.6e-127;
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 31 TEAEEDVDPWTVO---TSSAKCIDYDKLIVRGSSKIDKELINRIERATGAPPHFLR 86
DB 13 TESEOKITPEWVGAVVDGKSMGIDYDKLISOPGTGTHTEETLERKQVYGEHPHFLK 72

CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade.
 CC Note: This patent is in the same patent family as WO200065062, which
 CC contains additional sequences as shown in AAB99122-AAB99143,
 CC AAH75903-AAH75920 and AAG66436.

XX Sequence 385 AA;

Query Match 35.8%; Score 803; DB 22; Length 385;
 Best Local Similarity 45.3%; Pred. No. 1.3e-73;
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 35 EDFVDPVQVTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIFFSH 93
 DB 3 EDFVDPVQVTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIFFSH 58
 QY 94 RDMNQVIDAYENKPFYLYTGRGSSSEAMHGHILPFTFTWLODVNPVLYIOMTDEK 153
 DB 59 RDYDKVLQDYEEGRGFPLYTGRGPSG-PMHIGHIIPFATKWLQEKGVNLYIOTDDEK 117
 QY 154 YLMKD-LTLDOAYDAVENAKDIACGFDIKFTFISDLDYMGSSGFYKVKVRIQKHT 212
 DB 118 FLFENLTFEPTKHWATENIIDIAVGFDPPKTIPONSER---TKIYEMATIAKIN 173
 QY 213 FNOVKGIFGFTSDICIGISFPATQAAFSNSPQIFRRTDIQCLIPCAIDDPYFRM 272
 DB 174 FSKAKAVGFTQSGIKGIFFPATQIAPTF-----FEKR---RCLIPAIIDDPYWR 223
 QY 273 TRDVAPRIGYKPAALHSTFPALQCAOTKMSASDPNSIPLTPTAKOIKTKVKNHAFSG 332
 DB 224 QRDFAESLIGYKTAIHSKFVPSLTSCKMSASKPETATILTDSPEDEKVKWKFALTG 283
 QY 333 GRDIEEHROFGGCDVDSFMYLTFPLEDDDKLEQIRKDY---TSGAMLTEGLKALIE 389
 DB 284 GRPLKQREKGEPEKCVFKMLETFEEDDK--KIKERYACKNGELTGCCKRILIS 341
 QY 390 VLQPLIAHQARRKEVTDIYK 411
 DB 342 KIOEFLKEHQRRKKAEKQIEK 363

RESULT 15
 AAB58219
 ID AAB58219 standard; Protein; 142 AA.

XX AAB58219;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 557.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;

KW cardioactive; immunomodulatory; muscular active; vulnerary;

KM gastrointestinal; nephrotropic; antiinfective; gynecological;

KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

XX proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN WO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05918.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

DR N-PSDB; AAF18095.
 XX Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 XX
 PS Claim 11, Page 1051; 1425pp; English.

CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.

XX Sequence 142 AA;

Query Match 15.9%; Score 356.5; DB 21; Length 142;
 Best Local Similarity 78.0%; Pred. No. 2.3e-28;
 Matches 71; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 MSYAAAGDEYKADCPGPNAPPTSNHGPDTAEEDVDPVQVTSAAKIDYDKLIVR 60
 DB 52 MSYAAAGDEYKADCPGPNAPPTSNHGPDTAEEDVDPVQVTSAAKIDYDKLIVR 111
 QY 61 GSKIDKELINRIERATGQRPHFLRGIF 91
 DB 112 GSSXNQEELXDXP-STAKXTHS--GGGXFF 139

Search completed: October 24, 2002, 12:51:14
 Job time : 27.5607 secs

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.


```
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 58.1%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 4,1e-125;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

```
QY 29 DATEE-EDVDPTVOTSSAKGIDYKLYRFGSSKIDKELINRIRAIQORPHFLR 86
DB 7 DERESSEBQVNVPEVSAKDGKIDYKLDKFGCORLDESLLDRVORLTSRQPHVELR 66
QY 87 RGIFSRHDMNOVLDAVENKKPFYLYTGRGSSAMHGHILPIFTKWLQDVNPLVI 146
DB 67 RSVFRAHDFNEILDAYERGDKFTYITGRGSSSEALHGHILPIFTKYLQEAFFVPLVI 126
QY 147 QMTDEKYLKDLTDQAYGAVERNAKDIACGFDINTFTFLFSDIDYMGSSGFYKNVYK 206
DB 127 QLTDEKISIKTSLVEESQRLARENAKDIIACGFVTKTFTFSDIDYV--GAFYKNVYK 184
QY 207 IQKHTFNQVNGIRGFTSDICIGKISFPAIQAPSFNSFQIFEDRDIDQICLCAIDQ 266
DB 185 VGCVTLLKANGIRGFSSEDPKAKLSFPVQAVPSFPHLFPGKDNLCCLPCALDQ 244
QY 267 DPFYMTADVAPRIQYPPRIALHSTFPFALQOACTKMSASPNNSIFLTDRAKQIKTKVN 326
DB 245 DPFYMTADVAPRIQYSPALIESFFPALOGENGKMSASPNNSIYITDSKDKDKNNIN 304
QY 337 KHAESGSDTLEHRQFGSCNDVDSFMYLTFLEDDDKLEQIRKDYTSGMALGELKKA 386
DB 305 RYAESGGDSIEKHEKRELANLEVDIPVKYLSFFLEDDSELEHIKKEYEGCHMLTGEVKKR 364
QY 387 LIEVLOPLIAHQARRKVTDEIYEEFMTPRKLSFDQ 424
DB 365 LIEVLTIEVERHRRARAATVDEMADAFMAVRLPSKFE 402
```

RESULT 10
AAG23697
ID AAG23697 standard; Protein: 426 AA.
AC AAG23697;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27100.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
EN EP1033405-A2.
FN
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 14-JUN-1999; 9905-0138647.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.

PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144335.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 21-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 23-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.

QY 384 KKALEVLQPLIAEHQARRREVTDEIVKEMTPRKLSFDRQ 424
 DB 362 KKALEVLQPLIAEHQARRREVTDEIVKEMTPRKLSFDRQ 402

RESULT 6
 AAB47618 standard; Protein; 392 AA.
 XX AAB47618:
 AC 07-JAN-2002 (first entry)
 DT
 XX
 DE Human inactive TrpRS.
 XX
 KM Tyrosyl-tRNA synthetase; TyrRS; Rossmann fold nucleotide binding domain;
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
 KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KM dermal ulcer; diabetic ulcer; endothelialization;
 KM tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.
 XX
 OS Homo sapiens.
 XX WO200174841-A1.
 XX 11-OCT-2001.
 PD
 XX 21-MAR-2001; 2001MO-US08966.
 PF
 XX 31-MAR-2000; 2000US-193471P.
 PR
 XX (SCRI) SCRIPPS RES INST.
 PA Schimmel P, Wakasugi K;
 XX
 PI WPI: 2001-626377/72.
 XX N-PSDB; AAB43605.
 DR
 XX
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction -
 XX
 PS Disclosure; Page 135-136; 150pp; English.
 XX
 CC The sequences given in AAB47615-18 show full length and truncated
 CC versions of trptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
 CC the invention comprises a Rossmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic tRNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.
 XX
 SQ Sequence 392 AA:
 QY Query Match 88.5%; Score 1988; DB 22; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.8e-195.
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 47 SAKGIDYDKLIVRGSSKIDKELINRIERATGGRPHHFLRGJFFSHRDNQVLDAYENK 106
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 2 SAKGIDYDKLIVRGSSKIDKELINRIERATGGRPHHFLRGJFFSHRDNQVLDAYENK 61
 QY 107 KPFLYLTGRGSSPAMVGHILPITFTKWLQDVNVNPLVITQMTDDEKYLKMDLTDQAYG 166
 DB 62 KPFLYLTGRGSSPAMVGHILPITFTKWLQDVNVNPLVITQMTDDEKYLKMDLTDQAYG 121
 QY 167 DAVENAKDIACGFDINKTEFFSDLDYMGSSGFGYKNVYIKQKVFENQYKGIPTDSD 226
 DB 122 DAVENAKDIACGFDINKTEFFSDLDYMGSSGFGYKNVYIKQKVFENQYKGIPTDSD 181
 QY 227 CIGKISFPALQAPSEFSNSFPQIFRDRDITQCLIPCAIDDDPYFRMRYAPRIGYKPA 286
 DB 182 CIGKISFPALQAPSEFSNSFPQIFRDRDITQCLIPCAIDDDPYFRMRYAPRIGYKPA 241
 QY 287 LHSSTFPALQAGTQKASASPNSIFLTPTAKOIKTKVKNHAFSGGRDTEEHROFGN 346
 DB 242 LHSSTFPALQAGTQKASASPNSIFLTPTAKOIKTKVKNHAFSGGRDTEEHROFGN 301
 QY 347 CDVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALEVLQPLIAEHQARRREVT 406
 DB 302 CDVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALEVLQPLIAEHQARRREVT 361
 QY 407 DEYKERFMTPRKLSFDRQ 424
 DB 362 DEYKERFMTPRKLSFDRQ 379

RESULT 7
 ABB64621 standard; Protein; 430 AA.
 XX ABB64621:
 AC 26-MAR-2002 (first entry)
 DT
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 20655.
 XX
 KM Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 KM
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001MO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI: 2001-656860/75.
 XX N-PSDB; ABL08724.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL1840-ABL16175) and the encoded proteins
 CC (ABR57737-ABR72072).
 CC The sequence data for this patent did not form part of the printed

DR N-PSDB: AAX33942.
XX
PT New isolated human genes
XX
PS Claim 3; Page 112-114; 184pp; English.
XX
CC This sequence is encoded by a human gene of the invention, and is induced
CC to express by both HCMV and interferon (IFN), designated HCMV-inducible
CC genes (cig or cigs). The invention also relates to genes that are
CC repressed in the presence of HCMV infection, designated HCMV-repressible
CC genes (crg or crgs). The products can be used to obtain agents which can
CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
CC also be used for the development of drugs that would allow for higher
CC dosage IFN treatments without the concomitant toxicity normally
CC associated with administering high levels of IFN. The products can also
CC be used for detection, diagnosis and drug screening.
XX
SQ Sequence 471 AA;
Query Match 99.1%; Score 2226; DB 20; Length 471;
Best Local Similarity 99.3%; Pred. No. 7.9e-220;
Matches 421; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MSYKAGEDYKADCPGPNAPTNSHGPDATEAEDEVDPTVQTSSAKGIDYDKLIVRF 60
DB 48 MSYKAAAGEDYKADCPGPNAPTNSHGPDATEAEDEVDPTVQTSSAKGIDYDKLIVRF 107
QY 61 GSSKIDKELINIERATGQRPHHFLRGIFFSHRDMNOYLAVENKKPFYLYTGRGPSE 120
DB 108 GSSKIDKELINIERATGQRPHHFLRGIFFSHRDMNOYLAVENKKPFYLYTGRGPSE 167
QY 121 AMHVGHLPIETFTKWLQOVFNPLVIOGTMDEKYLMDLTLDOAYGDAVENAKDIIACGF 180
DB 168 AMHVGHLPIETFTKWLQOVFNPLVIOGTMDEKYLMDLTLDOAYGDAVENAKDIIACGF 227
QY 181 DINKTFITSDLDYMGSSGFYKNNVYKIQKHTFNOVKGIFGFTSDCIIGKISFPAIQAP 240
DB 228 DINKTFITSDLDYMGSSGFYKNNVYKIQKHTFNOVKGIFGFTSDCIIGKISFPAIQAP 287
QY 241 SFSNFPQIFPRRTDIOCLIPCAIDODPYFPRMTRVAPRIGYPRKALHSTFFPALQAG 300
DB 288 SFSNFPQIFPRRTDIOCLIPCAIDODPYFPRMTRVAPRIGYPRKALHSTFFPALQAG 347
QY 301 TKMSASDPNSSIFLTDTAQKQIKTKVKNHAFSGGRDIEEHROFGNCDDVYSFMYLTFFFL 360
DB 348 TKMSASDPNSSIFLTDTAQKQIKTKVKNHAFSGGRDIEEHROFGNCDDVYSFMYLTFFFL 407
QY 361 EDDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVTDIYKEFMTPRKLS 420
DB 408 EDDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVTDIYKEFMTPRKLS 467
QY 421 FDFQ 424
DB 468 FDFQ 471

RESULT 5

AAB47617
ID AAB47617 standard; Protein: 415 AA.AAB47617;
AC07-JAN-2002 (first entry)
DTHuman supermini TrpRS.
DE

XX Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW angiogenesis; graft; amycocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.
KW

X

OS Homo sapiens.
XX
PN WO200174841-A1.
XX
PD 11-OCT-2001.
XX
PF 21-MAR-2001; 2001WO-US08966.
XX
PR 31-MAR-2000; 2000US-193471P.
XX
PI (SCRI) SCRIPPS RES INST.
XX
PI Schimmel P, Wakasugi K;
XX
DR WPI: 2001-626377/72.
DR
DR N-PSDB: AHA43604.
XX
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
PS
PS Disclosure; Page 129-30; 150pp; English.
XX
CC The sequences given in AAB47615-18 show full length and truncated
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kDa molecular weight and is produced by cleavage of
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic RNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.
XX
SQ Sequence 415 AA;
Query Match 94.2%; Score 2116; DB 22; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.3e-208;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 SNHGPDATAEDEFDVDPVTQTSKAGIDYDKLIVRFSSKIDKELINIERATGQRPHH 83
DB 2 SNHGPDATAEDEFDVDPVTQTSKAGIDYDKLIVRFSSKIDKELINIERATGQRPHH 61
QY 84 FLRGIFFSHRDMNOYLDAVENKKPFYLYTGRGPSSSEAMHVGHLPIFTFKWLQOVNVP 143
DB 62 FLRGIFFSHRDMNOYLDAVENKKPFYLYTGRGPSSSEAMHVGHLPIFTFKWLQOVNVP 121
QY 144 LVIOGTMDEKYLMDLTLDOAYGDAVENAKDIIACGPIINTFTFSDLDYMGSSGFYKN 203
DB 122 LVIOGTMDEKYLMDLTLDOAYGDAVENAKDIIACGPIINTFTFSDLDYMGSSGFYKN 181
QY 204 VKIQRKHTFNOVKGIFGFTSDSDCIIGKISFPAIQAPFSNSFPQIFDRRTDIOCLIPCA 263
DB 182 VKIQRKHTFNOVKGIFGFTSDSDCIIGKISFPAIQAPFSNSFPQIFDRRTDIOCLIPCA 241
QY 264 IDDDPYFPRMTRVAPRIGYPRKALHSTFFPALQAGOTKMSADPNSSIFLTDRAKQIKT 323
DB 242 IDDDPYFPRMTRVAPRIGYPRKALHSTFFPALQAGOTKMSADPNSSIFLTDRAKQIKT 301
QY 324 KVNKHAFFSGGSDTIEEHROFGNCDDVYSFMYLTFFLEDDDKLQIRKDYSGAMLTGEL 383
DB 302 KVNKHAFFSGGSDTIEEHROFGNCDDVYSFMYLTFFLEDDDKLQIRKDYSGAMLTGEL 361

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
PS
XX Disclosure: Page 123-24; 150pp; English.
XX
CC The sequences given in AAB47615-18 show full length and truncated
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kDa Dalton molecular weight and is produced by cleavage of
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.
XX
SQ Sequence 437 AA;
Query Match 100.0%; Score 2246; DB 22; Length 437;
Best Local Similarity 100.0%; Pred. No. 6.2e-222;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
YY 1 MSYRAAGEDYKADCPNPAPTSNHGPDATAEEDPDPMTVOTSSAKGIDYDKLYRF 60
DB 1 MSYRAAGEDYKADCPNPAPTSNHGPDATAEEDPDPMTVOTSSAKGIDYDKLYRF 60
YY 61 GSSKIDKELINRIERATGQRPHHFLRRCIFFSHRDMNOVLDAENKKRPFLYTGSPSE 120
DB 61 GSSKIDKELINRIERATGQRPHHFLRRCIFFSHRDMNOVLDAENKKRPFLYTGSPSE 120
YY 121 AMHVGHLIPFIETKWLQOVFNVPVLTQMTDDEKYLMDLTLDQAYGDAVENAKDIACGF 180
DB 121 AMHVGHLIPFIETKWLQOVFNVPVLTQMTDDEKYLMDLTLDQAYGDAVENAKDIACGF 180
YY 181 DINKTIFSDLDYMGSSGFKYKVVYKIOKHTFNQVKGIFGFTSDCIGKISFPAIOAAP 240
DB 181 DINKTIFSDLDYMGSSGFKYKVVYKIOKHTFNQVKGIFGFTSDCIGKISFPAIOAAP 240
YY 241 SFSNSFPQIFRDRDIOCLIPCAIDQDPPYFMTQDVABRIGYKPPALLHSTFFPALQGAQ 300
DB 241 SFSNSFPQIFRDRDIOCLIPCAIDQDPPYFMTQDVABRIGYKPPALLHSTFFPALQGAQ 300
YY 301 TKMSASDNNSTFLLDTAKQIKTKYKNAHFGSGRDTIEHHQFGNCVDSFMYLTFEL 360
DB 301 TKMSASDNNSTFLLDTAKQIKTKYKNAHFGSGRDTIEHHQFGNCVDSFMYLTFEL 360
YY 361 EDDDLKLEIRKDYTSGAMLTGELKALIEVLOPLAEHQARKKVTDEIVKFEPMPTKLS 420
DB 361 EDDDLKLEIRKDYTSGAMLTGELKALIEVLOPLAEHQARKKVTDEIVKFEPMPTKLS 420
YY 421 FDFQ 424
DB 421 FDFQ 424
RESULT 2
AAB47615
ID AAB47615 standard; Protein: 484 AA.
XX
AC AAB47615;
XX
DT 07-JAN-2002 (first entry)
XX

DE Human full-length TrpRS.
XX
XX Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW triphophanyl-tRNA synthetase; trpRS; vascular graft surgery.
XX
XX Homo sapiens.
XX
XX WO200174841-A1.
XX
XX 11-OCT-2001.
XX
XX 21-MAR-2001; 2001MO-US08966.
XX
XX 31-MAR-2000; 2000US-193471P.
XX
XX (SCRI) SCRIPPS RES INST.
XX
XX Schimmel P, Wakasugi K;
XX
XX WPI: 2001-626377/72.
XX
XX N-PSDB; AAB43602.
XX
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
PS
XX Disclosure: Page 117-19; 150pp; English.
XX
CC The sequences given in AAB47615-18 show full length and truncated
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kDa Dalton molecular weight and is produced by cleavage of
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.
XX
SQ Sequence 484 AA;
Query Match 100.0%; Score 2246; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 7.2e-222;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
YY 1 MSYRAAGEDYKADCPNPAPTSNHGPDATAEEDPDPMTVOTSSAKGIDYDKLYRF 60
DB 48 MSYRAAGEDYKADCPNPAPTSNHGPDATAEEDPDPMTVOTSSAKGIDYDKLYRF 107
YY 61 GSSKIDKELINRIERATGQRPHHFLRRCIFFSHRDMNOVLDAENKKRPFLYTGSPSE 120
DB 108 GSSKIDKELINRIERATGQRPHHFLRRCIFFSHRDMNOVLDAENKKRPFLYTGSPSE 167
YY 121 AMHVGHLIPFIETKWLQOVFNVPVLTQMTDDEKYLMDLTLDQAYGDAVENAKDIACGF 180
DB 168 AMHVGHLIPFIETKWLQOVFNVPVLTQMTDDEKYLMDLTLDQAYGDAVENAKDIACGF 227
YY 181 DINKTIFSDLDYMGSSGFKYKVVYKIOKHTFNQVKGIFGFTSDCIGKISFPAIOAAP 240
DB 228 DINKTIFSDLDYMGSSGFKYKVVYKIOKHTFNQVKGIFGFTSDCIGKISFPAIOAAP 287

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 26.5607 Seconds
(without alignments)
1773.117 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSYKAAAGEDYKADCPGPNF.....VTDEIVKEFMTPRKLSDFDQ 424

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2246	100.0	437	22	AA847616 Human mini TyPRS.
2	2246	100.0	484	22	AA847615 Human full-length
3	2231	99.3	475	21	AA858220 Lung cancer associ
4	2226	99.1	471	20	AAV05372 Human HCMV inducib
5	2116	94.2	415	22	AA847617 Human supermini Tr
6	1988	88.5	392	22	AA847618 Human inactive Trp
7	1378	61.4	430	22	AB864621 Drosophila melanog
8	1378	61.4	430	22	AB867203 Drosophila melanog
9	1304	58.1	426	21	AA833698 Arabidopsis thalia
10	1304	58.1	426	21	AA833697 Arabidopsis thalia
11	1218.5	54.3	424	22	AA866931 Tryptophanyl-cRNA

12	968	43.1	292	21	AA823699 Arabidopsis thalia
13	831	37.0	173	21	AA858517 Lung cancer associ
14	803	35.8	385	22	AA896409 Putative P. abyssal
15	356.5	15.9	142	21	AA858219 Lung cancer associ
16	185.5	8.3	341	19	AA856423 Tryptophanyl tRNA
17	185.5	8.3	341	22	AA837851 Streptococcus pneu
18	164.5	7.3	409	20	AA894248 C. albicans tyrosy
19	162.5	7.2	409	20	AA894247 C. albicans tyrosy
20	159.5	7.1	379	20	AA837623 Protein involved i
21	158.5	7.1	378	22	AA896600 Putative P. abyssal
22	152	6.8	344	20	AA835439 Chlamydia pneumoni
23	146	6.5	339	22	AA835829 Helicobacter pylori
24	138.5	6.2	140	21	AA823032 Arabidopsis thalia
25	134.5	6.0	525	22	AA860745 Drosophila melanog
26	134	6.0	372	22	AA847614 Human mini TyPRS m
27	132.5	5.9	372	22	AA847612 Human mini TyPRS
28	132.5	5.9	536	22	AA847611 Human full length
29	128	5.7	334	22	AA835477 Haemophilus influe
30	123	5.5	347	21	AA831904 Arabidopsis thalia
31	122.5	5.5	346	21	AA848617 Arabidopsis thalia
32	120.5	5.4	404	22	AA854020 Propionibacterium
33	115.5	5.1	129	21	AA823033 Arabidopsis thalia
34	113.5	5.1	418	22	AA802578 Enterococcus tyros
35	113.5	5.1	423	22	AA835085 Enterococcus faeca
36	113	5.0	415	22	AA833998 Staphylococcus aur
37	113	5.0	420	18	AA819781 Tyrosyl-tRNA synth
38	113	5.0	420	22	AA837056 Staphylococcus aur
39	112	5.0	345	22	AA890498 C glutamicum prote
40	108	4.8	344	22	AA838213 Salmonella typhi c
41	107	4.8	71	22	AA810441 Human polypeptide
42	106.5	4.7	334	22	AA834778 E. coli cellular p
43	104.5	4.7	370	20	AA817508 pneumocystis carin
44	100.5	4.5	421	22	AA882363 S. epidermidis ope
45	95	4.2	418	18	AA827663 Streptococcus pneu

ALIGNMENTS

RESULT 1	AA847616 standard; Protein: 437 AA.
XX	AA847616;
AC	07-JAN-2002 (first entry)
XX	
DE	Human mini TyPRS.
XX	
KW	Tyrosyl-tRNA synthetase; TyPRS: Rossmann fold nucleotide binding domain;
KW	vascular endothelial cell function; burn; plastic surgery; abdomen;
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW	dermal ulcer; diabetic ulcer; endothelialization;
KW	tryptophanyl-tRNA synthetase; tyPRS; vascular graft surgery.
XX	
OS	Homo sapiens.
XX	
PN	WO200174841-A1.
XX	
PD	11-OCT-2001.
XX	
PF	21-MAR-2001; 2001MO-US08966.
XX	
PR	31-MAR-2000; 2000US-193471P.
XX	
PA	(SCRI) SCRIIPS RES INST.
XX	
PI	Schimmel P, Wakasugi K;
XX	
DR	WPI; 2001-626377/2.
XX	N-PDB: AA843603.

RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000007; BAA31046.1; .
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 KW Aminoacyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 301 AA; 35312 MW; 957143EE1647CAE1 CRC64;

 Query Match 24.6%; Score 626.5; DB 17; Length 301;
 Best Local Similarity 44.2%; Pred. No. 2e-45; Mismatches 92; Indels 21; Gaps 7;
 Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

 QY 169 MHVGLIPEITFKMLQDVENVPLVIOMTDEKYLKMD-LTLQAYGDAVENAKDIACGF 227
 DB 1 MHGHIIPFATKMLQEKGVNLVYIOTDDEKFLFRENLTFFDTRKMAVDNLIDIAVGF 60
 QY 228 DINKTFIESDLYMGMSGCFYKNVYKIQKHVFNQVKGIFGTFSDCICKISFPALQAP 287
 DB 61 DDKTFEIPQNSF---TKIYEMAIPIARKINFSMAKAVGFEQSGIKGIMIFFPAILQAP 116
 QY 288 SFSNSFPQIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFEPALQAG 347
 DB 117 TF-----FEKK---KCLIPALIDDDPYWRLORDPAESLGYYKTRALHSKFVPSLTSL 166
 QY 348 TMSASDPNSSIFLDTAKQIKTKYKNAFSGGRDTEIEHRQFGNCVDVSEFMVLTFFL 407
 DB 167 GKMSASKEPTAIYLDSPEDVEKWKFTLTGSRPLTKQREKCGCECVFKNWEIIF 226
 QY 408 EDDDLKLEQIRKY---TSCAMLTGELKKALIEVLOPLIAEHQARRKEVTDEIVKEFMP 464
 DB 227 EDDDK--KLKERYACKNELTCGECRKLISKIOEFLKEHQRRRK-AEKLVKRYKYG 283
 QY 465 KLS 467
 DB 284 KLA 286

 RESULT 15
 ID Q9HN66 PRELIMINARY; PRT: 380 AA.
 AC Q9HN66; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TRYPTOPHANYL-TRNA SYNTHETASE.
 GN TRPS2 OR VNG2232G.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroga J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angeline C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
 RT "Genome sequence of Halobacterium species NRC-1."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE005109; AAG20355.1; .
 DR InterPro: IPR001994; Cytidylyltransf.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; UNKNOWN_1.
 KW Aminoacyl-tRNA synthetase; Complete proteome.

SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;
 Query Match 21.0%; Score 535.5; DB 17; Length 380;
 Best Local Similarity 35.2%; Pred. No. 1.8e-37;
 Matches 142; Conservative 78; Mismatches 142; Indels 41; Gaps 17;

 QY 80 AEEDEVDPWTYOTSSAKGIDYDKLLVRFSSKIDELINRTERATGQRPNHLRGIFFS 139
 DB 3 ADGNQVTPYAVESDD--LDYEKLLARFGADELTDQDARFP-----DHPLVNRGLFYA 53
 QY 140 HEDMNOUVDAYENKRPFLYITGRGSSSEAMVGHILPFTKWLQDVENVPLVIOMTDE 199
 DB 54 GRDQVDFPLAGQS---IVTGVGPSG--PMHIGHAMVEFYFARRLQDEFGARYVPLSDE 108
 QY 200 KYLMVDLTQAYGDAVE-NAKDIACGFDINKTFIF---SDLDYV-GMSSGFYKNYKI 254
 DB 109 KYWFDQTPAET-GDYLKRLNLDLAVGFDPELTIYVDTRADVLYPLATFAGDV--- 164
 QY 255 OKHVFENVKGIFFGTFSDCICKISFPALQAPSPSNSFPQIFRDRTDIOCLIPCAIDOD 314
 DB 165 -RHATLQNVYG-----EPDNGQAFYPAVQTAHL--LPQLVHG--EHETLVPIAVDOD 213
 QY 315 PYFRMTRDVAPRIGYP--KPALHSTFEPALQAGQOTKMSASDPNSSIFLDTAKQIKTKV 372
 DB 214 PIVRVSBDVAAKARYPVGKPGALLMQFLPSLAG-PGKMSSS-AGVSIRLTSPDTRREKV 271
 QY 373 NKHAFSGGRDTEIEHRQFGNCVDVSEFMVLTFFLEDD-KLEQIRKDYTSQAMLTGELK 431
 DB 272 RTHATYTGKASVEEHRKAKGVPAEDVPFOYLSAFEPDDAELARKEREYRAGDLSGELK 331
 QY 432 KALIEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFOKLA 474
 DB 332 DLAADRTFEFLAAHQRRRALGD--VTEALDAFRLTDERQRA 372

Search completed: October 24, 2002, 12:53:03
 Job time : 31.2521 secs

us-09-813-718-10.rspt

QY 427 TGE LKALIE-VLØPLAEHQARKEVTDEIVKEEMTPRKL 466
::: : :: : : : : : : : : : : :
Db 448 SGEVKDLAADCIITREVLHMDERRRTVTDDDVEECRINDI 488

RESULT 13
Q95295

ID	PRELIMINARY:	PRT:	136 AA.
AC	Q95295;		
DT	01-FEB-1997 (TrEMBLrel. 02, Created)		
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	TRYPHOPHANYL-TRNA SYNTHASE (FRAGMENT).		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=SMALL INTESTINE;		
RA	Winteroe A.K., Fredholm M., Davies W.;		
RT	"Evaluation and characterization of a porcine small intestine cDNA		
ST	library.";		

RL	Submitted (JUL-1995)	to the EMBL/GenBank/DBD databases.
DR	EMBL: Z81267; CAB03585.1; -	
DR	InterPro: IPR002306; tRNA-synt_tlp.	
DR	PRINTS: PR01039; tRNAsyntHTRP.	
FT	NON_TER	1
FT	NON_TER	136
SO	SEQUENCE	136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;
Query Match	26.8%;	Score 682; DB 6; Length 136;
Best local Similarity	94.1%;	Pred. No. 1.1e-50;
Matches 128;	Conservative	2; Mismatches 6; Indels 0; Gaps
QY	281 PAIOAAEFNSFQIIFRDRTDIOCLIPCALDOPFRMTADVAPRIGYKRPALHSTFF	340
Db	1 PAIQABPSFSSFPQIRDRDRIODCLIPCALDOPFRMTADVAPRIGYKRPALHSTFF	60
QY	341 PALGGAQTCKMSASDPNSSIFLTDATAKIKTVKNNAHSGGRDITIEHRQFGGNCDDVST	400

Db 61 PALGGAQTKMSXXAPNSSIF
QY 401 MYLTFLEDDDKLEQI 416

Db 121 MYLTFLEDDRXEQI 136

RESULT 14
059584
ID 050594
DDEBTMTNARY.

ID	059584	PRELIMINARY;
AC	059584;	
DT	01-AUG-1998	(TREMBLrel. 0
DT	01-AUG-1998	(TREMBLrel. 0

DT 01-JUN-2001 (TREMBLER, 1
DE 301AA LONG HYPOTHETICAL T
GN PH1921.

OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; T
OX NCBI_TaxID=53953;

RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=OT3;

RX MEDLINE=98344137; PubMed=
RA Kawarabayasi Y., Sawada M
RA Yamamoto S., Sekine M., B

RA Sakai M., Ogura K., Otsuk
RA Funahashi T., Tanaka T.,
RA Aoki K.-I., Yoshizawa T.,

RA Masuchi Y., Shizuya H., K
RT "Complete sequence and ge
RT thermophilic archaeobacter
. RT

RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.";

[illegible]

RESULT 10			
ID	Q976M1	PRELIMINARY;	PRT; 406 AA.
AC	Q976M1;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, last annotation update)		
DE	PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.		
GN	ST0169.		
OS	Sulfolobus tokodaii.		
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.		
OX	NCBI_TaxID=111955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=JCM 10545 / ?;		
RX	PubMed=11572479;		
RA	Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y., Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.;		
RA	"Complete genome sequence of an aerobic thermocacidophilic Crenarchaeon, Sulfolobus tokodaii strain7."		
RT	DNA Res. 8:123-140(2001).		
RL	EMBL: AP000981; BAB65126.1; -.		
KW	Aminacyl-tRNA synthetase; Hypothetical protein; Complete proteome.		
SO	SEQUENCE 406 AA; 47979 MW; 276CCD6506BC94B3 CRC64;		

Query Match	35.7%;	Score 910.5;	DB 17;	Length 406;
Best Local Similarity	50.5%;	Pred. No. 1.5e-69;		
Matches 189;	Conservative 63;	Mismatches 99;	Indels 23;	Gaps 9;

```

QY      82 EDF -VDPMTTVOGSSAKG -IDYKTLVRFSSKSIDELNIRIRATGQRPHHLRKRFPS 139
QY      28 QDENFTPHEV-----KGVNDYDKLLVQFSTQKTIITSELKIKISIIDELHAWLRKRVFSS 82
QY      140 HRDMQVLDAYENKRPFLYTGSGSSSEAMVGHILPIETFKWLDQFVNPLVIGQMTDE 199
QY      83 HRDLDLVLEKDYDQDGKGFELTYGKRAPL -GMHIGHLPIETFKWLDQKRFVNLYIETTDE 141
QY      200 KYLWK -DLTDQAYDADAYENAKDITACGFDINKTIFESLDYMGSSSGYKRVKVIQKIV 258
QY      142 KEMRNPETYLDTQROMAWADNIDIIITAVGPNPKETFIQDEYI---RNNYPAIRAKIAKKL 198
QY      259 TFENQKJLGFQFSDSCIGKISFPALQAAPSFNSNPQJFRODRTDIOQLCPALDQDPYR 318
QY      199 TFSEVRATFEGDLTSSNIGITWIPALQIAAT-----MEKR---RCLIPAGIDQDPYR 248

```

[illegible]

RESULT 11

Q97ZX0		
ID	097ZX0	PRELIMINARY;
AC	097ZX0	PRT;
DT	01-OCT-2001 (TReMBLrel. 18, Created)	386 AA.
DT	01-OCT-2001 (TReMBLrel. 18, Last sequence update)	
DT	01-OCT-2001 (TReMBLrel. 18, Last annotation update)	
DE	TRYPTOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).	
GN	TRPS.	
OS	Sulfolobus solfataricus.	
OC	Archaea: Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.	
OX	NCBI_TaxID=2287;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 35092 / DSM 1617 / P2;	
RX	MEDLINE=11337296; PubMed=1142776;	
RA	She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,	
RA	Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,	
RA	De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,	
RA	Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,	
RA	Thi-Noc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,	
RA	Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,	
RA	Garrett R.A., Regan M.A., Sensen C.W., Van der Oost J.;	
RT	"The complete genome of the crenarchaeon Sulfolobus solfataricus P2."	
RL	Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).	
DR	EMBL; AE006677; AAK40778.1; -.	
DR	InterPro: IPR002305; tRNA-synt.1b.	
DR	InterPro: IPR002306; tRNA-synt.1rp.	
DR	Pfam: PF00579; tRNA-synt.1b; 1.	
DR	PRINTS; PR01039; TRNASYNTHTRP.	
QW	Aminoacyl-tRNA synthetase, Lysine, Complete proteome.	
QW	SEQUENCE 386 AA; 45448 MW; 3F4DA8C9B8D219DF CRC64;	

Query Match	35.6%	Score 907	DB 17	Length 386
Best Local Similarity	48.6%	Pred. No. 2.8e-69		
Matches 191	Conservative 68	Mismatches 110	Indels 24	Gaps 10

QY	78	TEAEDEFADPTVOTSSAKG - IDDKLTVREVGSSKIDKELINRERATGCRPHHPLRRI	136
Db	6	TMPEDEFTYTPREv-----KGNVDKDKLIVGPGTOKITEELKOKRIKKNLAGSL - HVMLRRNV	59
QY	137	FFSHRDANOVLDAVENKPKPLYLYTGRGPSSBAHVGLPLPIFTKWLQDVFNPLVLYQMT	196
Db	60	FFSHRDIDLVLNDYKSGFELLYTGRAPSL - GMHIGHLPIFTFKWLOEKFANLYIEIT	118
QY	197	DDEKYLWK - DLTLLOAGDAVENKDKITPACGDFINKEFISDDLYDMSSGFCFKNVVKIO	255
Db	119	DDEKYMNRPEFTLDQTSMAINDLNLINGVFNDRKFIYQDHEI---RNMYPIYVKA	175
QY	256	KHVTFNQVYKGIFFGTDTSDCIQIKISFPAIQAPSFSSNSFOIHFDRYDIOCLIPCAIDQD	315
Db	176	KKLTFSEVRAPFFGIDASSNIGLIFYPALQIAPT-----MPEKK--RCLIPAGIDQD	225
QY	316	YFRMTRDVAAPRIGYKPAALLHSTFEPALQOAKRMSADPNSSPLTDTKQIKTKTVKNH	375
Db	226	YWRLOQRIABESLYGKKAQOIKSHKLLPPLTPEBGKMSSNPFATILYVDKTYERKIMKY	285
QY	376	AFSGGRPTIEEHROFGNCDVDVSMYLTFLEDDD - KLEQIKRDYTSGAMLTELKALKAL	434
Db	286	AFSGGQPTIELHRRYGGNPBLVDVFPQWLYYFEDDDNRIRKEIEEYRSGMILTELKOLLI	345

"The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).
 DR EMBL: AE003682; AAC22136.1; -
 DR FlyBase: FBgn0010803; Aats-trp.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 DR SEQUENCE 430 AA; 47971 MW; 2938EECC69E979F CRC64;

Query Match 54.1%; Score 1378; DB 5; Length 430;
 Best Local Similarity 61.4%; Pred. No. 1.5e-109;

Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

OY 57 DYKADCPGPNPAITSNHCGBDA-----TEAE-----EDFVDPWTYQTSASAKIDYD 101
 DB 3 DKEKEIVEGEVLEALINGKRDADPEVETGDDAQEGATAPTEDVDVPMWVASSNDAGVDYD 62
 OY 102 KLIVRFGSSKIDKELINRERATGPRPHHFLRGIFESHROMNQVLDAAVENKKPPYLITG 161
 DB 63 KLIRKFGSSKIDKELIAREKITGKPAHFFIRGMPFSHRDLHTLTTLREGCKPPYLITG 122
 OY 162 RGPSEAMHVGHLIPFIFTKLQDVFNPLVQMTDDEKYLKDLTDQAYGDAVENAKD 221
 DB 123 RGPSSGSLHVGHLVPEFIMTKWLOEFEDVPLVQLTDDEKTLKDLKVEDAIKLGRENKD 182
 OY 222 ITACGPDINKTFIFSDLDYMGSSGFEYKNVAKIQKHVTFENQVKGJFEGTDSICIGISFP 281
 DB 183 IVAIFEDVKNFTFIFNNLEFEVCKPAMYONILRIQKCVTFENQVKGJFEGTDSIIKIGRP 242
 OY 282 AIQAAPDSNSPQPIFRORTDIOCLIPGALDODPYFRMRDVAAPRGVAPKALLHSTFP 341
 DB 243 AAQAAPALSTSTPFLFGN-KVHCLIPGALDODPYFRMRDVAAPRGVAPKALLHSTFP 301
 OY 342 ALOGAQTKASDPSNIFLDTAKQIKTKVKNHAFSGGRDTEEHROFGNCDVDVSPM 401
 DB 302 ALQGAKTMSASDQNSAVYLTDTPKQIKKINKYAFSGSRVSVFEHRKRGKGPEDVSYQ 361
 OY 402 YLTFPLEDDDKLEQIRKDYTSGLMTGELKALILEVLOPLAHOARRKEVTDEIVKEFM 461
 DB 362 LKFELEDADAKLEEVAVYVSKGEMLTGELIKLAVETLPIVBOHOARRKLLIDVELDKYF 421
 OY 462 TPKRLSF 468
 DB 422 ELRPLKF 428

RESULT 8

O9SR15 PRELIMINARY: PRT; 402 AA.

AC O9SR15
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUVATIVE TRYPTOPHANYL-TRNA SYNTHETASE.
 GN F7018.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Beilto M., Creasy T.H., Haas B.,
 RA Rongning C.M., Koo H., Fujii C.Y., Uteback T.R., Barstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RA "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC011437; AAF04890.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.

DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75EE56CD15 CRC64;

Query Match 51.1%; Score 1301; DB 10; Length 402;
 Best Local Similarity 60.3%; Pred. No. 5.3e-103;
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

OY 76 DATEAE--EDFVDPWTYQTSASAKIDYDKLIYRFGSSKIDKELINRERATGPRPHHFLR 133
 DB 7 DEREAESSQVYVNPWEVSAKDGKIDYDKLIDKFCGQRDESLDRVQLRTRQGHVFLR 66
 OY 134 RGIFFSHRDMNOVLDAAVENKKPPYLITGGRPSSEAMHVGHLIPFIFTKLQDVFNPLVI 193
 DB 67 RSVFPAHRDFNELLDAVERGDFLYLTGKPSSEALHGLHLPMTFTYLOEAFKVPPLVI 126
 OY 194 QMTDDEKYLKDLTDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFEYKNVAK 253
 DB 127 QLTDEKSTIKMLVSESOQLARENKADIIACGPDYTKTFIFSDPDYV--GAFYKNVAK 184
 OY 254 IQKHVTFENQVKGJFEGTDSICIGISFPALQAAPDSNSPQPIFRDRDIOCLIPCAIDQ 313
 DB 185 VCKCVTLKAMGIFEGSGEDPIAKLSFPVYQAVSPSPSPHLPFGKDWLRCLIPCAIDQ 244
 OY 314 DPFPMTRDVAAPRGVAPKALLHSTFPALQGAQTKMSASDPSNIFLDTAKQIKTKVN 373
 DB 245 DPFPMTRDVAAPRGVAPKALLHSTFPALQGAQTKMSASDPSNIFLDTAKQIKTKVN 304
 OY 374 KHAFFSGGRDTEEHROFGNCDVDVSPMFLPFLEDDDKLEIRDYSGAMLTELKKA 433
 DB 305 RAFFSGGQDSITKHRELGNLEVDLPVKYLSFLEDDSELEIKKEYGGRMLTEGVKKR 364
 OY 434 LIEVLOPLAHOARRKEVTDEIVKEFMTPKRLSFDFQ 471
 DB 365 LFEVLTEIVERRRARRAANTDEMVAFAFNAVRLPMPFE 402

RESULT 9

O9UIR2 PRELIMINARY: PRT; 324 AA.

AC O9UIR2
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Y80D3A.1 PROTEIN.
 GN Y80D3A.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B.R.;
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: AL132853; CAB60439.1; -
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 SQ SEQUENCE 324 AA; 36289 MW; 66687E6E420ECC12 CRC64;

Query Match 40.3%; Score 1027; DB 5; Length 324;

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
DR EMBL: AK007754; BAB25235.1; -.
DR MGI: 104630; Mavs.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
SO SEQUENCE 329 AA; 37613 MW; 701E702DC244CA2A CRC64;

Query Match 64.0%; Score 1629; DB 11; Length 329;
Best Local Similarity 92.1%; Pred. No. 3.2e-131;
Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 143 MOVDAVENKKPFYLYTGRGSSSEAMVHGLIPFTKWLQDVFNPLVLIQMTDDEKYL 202
DB 1 MNOUIDAENKKPFYLYTGRGSSSEAMVHGLIPFTKWLQDVFNPLVLIQMTDDEKYL 60
QY 203 WKDLTLQAYGDAVENAKDIACGFDINKFTFSDLDYMGSSGFYKVVKIQKHVTENQ 262
DB 61 WKDLTLQAYGDAVENAKDIACGFDINKFTFSDLEYMGSGFYRVNVKIQKHVTENQ 120
QY 263 VKGIFGFDSQICIGKISFPALQAAFSFNSPFOIFRDRTDIOCLIPCALIDDPYRMRD 322
DB 121 VKGIFGFDSQICIGKISFPALQAAFSFNSPFIKFRDRTDIOCLIPCALIDDPYRMRD 180
QY 323 VAPRIYKPKPALHSTFFPALQAGTQKMSASDPNSSIFLDTAKOIKTKVKNHAFSGRD 382
DB 181 VAPRIYKPKPALHSTFFPALQAGTQKMSASDPNSSIFLDTAKOIKTKVKNHAFSGRD 240
QY 383 TIEEHROFGNCDVDSFWYLFLEDDDKLEQIRKDYTSGLTGLKALIEVLOPLI 442
DB 241 TIEEHROFGNCDVDSFWYLFLEDDDKLEQIRKDYTSGLTGLKALIEVLOPLI 300
QY 443 AEHQARRKEVDELYKEFMTPRKLSFDQ 471
DB 301 AEHQARRKAVTEYKFEFMTPRQLSFHQ 329

RESULT 4
070184 PRELIMINARY; PRT; 305 AA.
AC 070184;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TRYPTOPHAN-TRNA SYNTHETASE (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathia; Cavidae; Cavia.
OX NCBI_Taxid=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HARTLEY; TISSUE-SPLEEN;
RA Yang D., Goto R., Watanabe N., Kobayashi Y.;
RT "Identification and Cloning of Genes Whose Expressions are Elevated
RT during DMCB-Induced Guinea Pig Skin Delayed-type Hypersensitivity
RT Reaction".
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB012222; BAA25288.1; -.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 1
SO SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 60.3%; Score 1537; DB 11; Length 305;
Best Local Similarity 94.4%; Pred. No. 2.2e-123;
Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 167 EAMVGHILIPFTKWLQDVFNPLVLIQMTDDEKYLKDLTLQAYGDAVENAKDIACG 226
DB 1 EAMVGHILIPFTKWLQDVFNPLVLIQMTDDEKYLKDLTLQAYGDAVENAKDIACG 60
QY 227 FDINKTFEFDLDYMGSSGFYKVVKIQKHVTENQVGIIGFTDSDICISFPALQAA 286
DB 61 FDINKTFEFDLEYMGSGFYKVVKIQKHVTENQVGIIGFTDSDICISFPALQAA 120
QY 287 PSFNSPFOIFRDRTDIOCLIPCALIDDPYRMRDVAIPRIYKPKPALHSTFFPALQGA 346
DB 121 PSFNSPFOIFRDRTDIOCLIPCALIDDPYRMRDVAIPRIYKPKPALHSTFFPALQGA 180
QY 347 QTKMSASDPNSSIFLDTAKOIKTKVKNHAFSGRDTIEEHROFGNCDVDSFWYLF 406
DB 181 QTKMSASDPNSSIFLDTSAKOIKTKVKNHAFSGRDTIEEHROFGNCDVDSFWYLF 240
QY 407 LEDDDKLEQIRKDYTSGLTGLKALIEVLOPLIAHQARRKEVDELYKEFMTPRKL 466
DB 241 LEDDDKLEQIRKDYTSGLTGLKALIEVLOPLIAHQARRKEVDELYKEFMTPRKL 300
QY 467 SFDFQ 471
DB 301 SFDFQ 305

RESULT 5
0904Y1 PRELIMINARY; PRT; 430 AA.
ID 0904Y1;
AC 0904Y1;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHETASE.
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99250164; PubMed=10233165;
RA Seshalah P., Andrew D.J.;
RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
RT the developing Drosophila salivary gland".
RL Mol. Biol. Cell 10:1595-1608(1999).
DR EMBL: AF125156; AAF20166.1; -.
DR FLYBase: FBgn010803; Aats-trp.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase.
SO SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;

Query Match 54.3%; Score 1382; DB 5; Length 430;
Best Local Similarity 61.6%; Pred. No. 6.8e-110;
Matches 263; Conservative 63; Mismatches 85; Indels 16; Gaps 3;

QY 57 DYKADCPGNPAPTSNMGPD-----TEAE-----EDFYDPMVTVQTSAGIDYD 101
DB 3 DTKETVVGCVATLTNGCPDAEPVETGDAQAQEGATPFDVDVPMVAVASSNDAGVYD 62
QY 102 KLIYFGSSKIDKELINIERATGQRPHPLRGIFGSHRDMNOVLQAYENKKPPYLYTG 161
DB 63 KLIKRFSSKIDDELIAFEKITGKPAHHFTRGWFESHRLTILTLREGKRPFLYVTG 122

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 26.2521 Seconds
(without alignments)
3189.442 Million cell updates/sec

Title: us-09-813-718-10

Perfect score: 2547
Sequence: 1 MPNSEPASLLEFNSIATQG.....KLSFDFQKLAALHHHHHH 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2227.5	87.5	475	11	Q9DC65	Q9dc65 mus musculus
2	2223.5	87.3	481	11	Q9J558	Q9j558 mus musculus
3	1629	64.0	329	11	Q9D8R9	Q9d8r9 mus musculus
4	1537	60.3	305	11	070184	070184 cavia porce
5	1382	54.3	430	5	Q9U4Y1	Q9u4y1 drosophila
6	1379.5	54.2	420	5	Q9U4Y0	Q9u4y0 drosophila
7	1378	54.1	430	5	Q9VHG2	Q9vhg2 drosophila
8	1301	51.1	402	10	Q9SR15	Q9sr15 arabidopsis
9	1027	40.3	324	5	Q9U1R2	Q9u1r2 caenorhabd
10	910.5	35.7	406	17	Q976M1	Q976m1 sulfolobus
11	907	35.6	386	17	Q97ZX0	Q97zx0 sulfolobus
12	800.5	31.4	490	5	Q9U1F5	Q9u1f5 leishmania
13	682	26.8	136	6	Q9S295	Q9s295 sus scrofa
14	626.5	24.6	301	17	Q9S584	Q9s584 pyrococcus
15	535.5	21.0	380	17	Q9H866	Q9h866 halobacteri
16	428	16.8	136	6	Q9T588	Q9t588 bos taurus

17	310	12.2	111	5	Q95YU8	Q95yl8 encephalito
18	299.5	11.8	426	17	Q978Y8	Q978y8 thermoplasma
19	294	11.5	157	5	Q9U533	Q9u533 trypanosoma
20	290	11.4	109	4	Q9UD15	Q9ud15 homo sapien
21	274.5	10.8	513	17	Q9HN83	Q9hn83 halobacteri
22	268.5	10.5	426	17	Q9H1W5	Q9h1w5 thermoplasma
23	188	7.4	364	17	Q9YA64	Q9ya64 aeropyrum p
24	185.5	7.3	341	16	Q97N42	Q97n42 streptococc
25	183	7.2	340	16	Q99XH4	Q99xh4 streptococc
26	181.5	7.1	331	17	Q97921	Q97921 thermoplasma
27	174.5	6.9	341	16	Q9CJDI	Q9cjd1 lactococcus
28	168.5	6.6	895	10	Q9SGN2	Q9sgn2 arabidopsis
29	167	6.6	351	16	Q9RVD6	Q9rvd6 deinococcus
30	163	6.4	460	10	P93018	P93018 arabidopsis
31	159	6.2	33	4	P78534	P78534 homo sapien
32	158.5	6.2	375	17	Q9V027	Q9v027 pyrococcus
33	155	6.1	1714	5	Q95T13	Q95t13 drosophila
34	154.5	6.1	102	1	Q07119	Q07119 halobacteri
35	152.5	6.0	682	5	Q9N9B8	Q9n9b8 leishmania
36	149.5	5.9	327	17	Q9HN62	Q9hn62 halobacteri
37	149.5	5.9	408	10	P93363	P93363 nicotiana t
38	147	5.8	528	11	Q91WQ3	Q91wq3 mus musculus
39	146	5.7	375	17	Q58739	Q58739 pyrococcus
40	145	5.7	525	5	Q9VW60	Q9vw60 drosophila
41	144	5.7	332	17	Q9HKT3	Q9hkt3 thermoplasma
42	140.5	5.5	1149	5	Q23315	Q23315 caenorhabd
43	139	5.5	294	17	Q96YV3	Q96yv3 sulfolobus
44	138.5	5.4	449	5	Q9NG35	Q9ng35 drosophila
45	138.5	5.4	449	5	Q9NG35	Q9ng35 drosophila

ALIGNMENTS

RESULT 1
ID Q9DC65 PRELIMINARY; PRT; 475 AA.
AC Q9DC65;
DT 01-JUN-2001 (TREMBLrel, 17, Created)
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DE ADULT MALE LUNG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
CLONE:1200002C07, FULL INSERT SEQUENCE.
GN WARS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
RA Blake J., Boffelli D., Bojunga N., Carlinici P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamly M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK004541; BAB23357.1; -.


```

AC 046127;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR TRSA.
OS Clostridium longisporum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1533;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6405;
RA Brown G.D., Thomson J.A.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: LA9336; AAC05711.1; -
DR HSSP: P00953; IDZR.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b: 1.
DR PRINTS: PR01039; TRNASYNTTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 12 20 "HIGH" REGION.
FT SITE 201 205 "KMSKS" REGION.
FT BINDING 204 204 ATP (BY SIMILARITY).
SQ SEQUENCE 341 AA; 38256 MW; 693C820F5A0844D1 CRC64;

Query Match 7.0%; Score 178.5; DB 1; Length 341;
Best Local Similarity 24.2%; Pred. No. 5.2e-07;
Matches 81; Conservative 60; Mismatches 159; Indels 35; Gaps 11;

OY 158 LYTGRGPSEAMHVGHLIPFTK-WLQDVENVPLVIOQMTDDEKYLKDLTLDAQAYGAV 216
DB 6 ILTGDRPTGK-LHIGHYVSLKRNVOLONSGDSRFSIMADQALTDNARNPEKIRNSLI 64
OY 217 ENAKKILICGDFINTKTFE--SDLDYMGSSGFYKNV---KIQKHVFENQKGIFFGD 271
DB 65 EYALDYLAAGDPLKSTLIVOSQPELNLNHYLNLVTLSELRNPYVKAIEKQKNFEN 124
OY 272 SDICIKISFPAIOAAPSFSNSFPQIFRDTIOCLIPCAIDDPYFRMTRDVAPRIG-- 328
DB 125 STPAGFLIYVPSQADITN-----FKATY-----VPVGEDQLPMIEQARELVRSGNTLY 173
OY 329 -----YKRPALLHSTF--FPALQAGQTKMSASDPNSSTFLTDYTKQKTKYKNAKIAFSGG 380
DB 174 GKEVLVERPAVIRPKGTIGRLPCTDG-KAKMSKISGN-ALVLADEADVIKQKWSMYTDPN 231
OY 381 RDTIEEHROFGNCVDVAFMYLTFELEDDEKLEQIRKQYTSGAMLTGELKALLEVLP 440
DB 232 NIKVTDPGOVEB-----YVFTYLDTFCKDTLETLEKKAHYSNGIGDYKVKVFLNELLDA 287
OY 441 LIAEQARKREYDEIVKEFMPTRKLSFDPFKIAA 475
DB 288 ELEPIRNRKKEFOKDIPREYIRLKSGSEKAREVAA 322

```

```

RESULT 15
SYCC_YEAST
ID SYCC_YEAST STANDARD: PRT: 394 AA.
AC P36421;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tyrosyl-tRNA synthetase, cytoplasmic (EC 6.1.1.1) (Tyrosyl--tRNA
DE ligase) (TYRS).
GN TYSL OR MGM104 OR YGR185C OR G7522.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93286133; Pubmed=8509419;
RA Chow C.M., Rajbhandary U.L.;
RT "Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene.
RT Isolation by complementation of a mutant Escherichia coli suppressor
RT tRNA defective in aminoacylation and sequence analysis.";
RL J. Biol. Chem. 268:12855-12863(1993).
RN [2]
RN SEQUENCE FROM N.A.
RA Guan M.-X., Chen X.-J., Clark-Walker G.D.;
RL Submitted (MAY-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; Pubmed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363(1997).
RN [4]
RP SEQUENCE OF 1-36 FROM N.A.
RC STRAIN=B3926;
RX MEDLINE=95087887; Pubmed=7995524;
RA Henry N.L., Campbell A.M., Feaver W.J., Poon D., Well P.A.,
RA Kornberg R.D.;
RT "TritF-TAF-RNA polymerase II connection.";
RL Genes Dev. 8:2868-2878(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L12221; AAB59329.1; -
DR EMBL: X71998; -; NOT ANNOTATED_CDS.
DR EMBL: 272970; CAA97211.1; -
DR EMBL: X99074; CAA67529.1; -
DR EMBL: U13015; AAB61641.1; -
DR PIR: A45999; A45999.
DR SCD: S0003417; TYSL.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002307; tRNA-synt_tyr.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01040; TRNASYNTTYR.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 48 56 "HIGH" REGION.
FT SITE 227 231 "KMSKS" REGION.
SQ SEQUENCE 394 AA; 44020 MW; 57EBDB9BE6D054B7 CRC64;

```

```
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000064; BAA81476.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 81 89 "HIGH" REGION.
FT SITE 258 262 "KMSKS" REGION.
SQ SEQUENCE 374 AA; 42400 MW; A72635B7CA3F9189 CRC64;

Query Match 13.9%; Score 353; DB 1; Length 374;
Best Local Similarity 30.4%; Pred. No. 5.2e-21;
Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 85 VDPMTVQVSSAKGIDYDKLIVRFSGSSKIDKELINIERATGQRPHHFLRGIFFSHRDMN 144
DB 8 LDPW---GAVEIKDYDRLRTFGIRPSEVL--PLRKAMEBPSFLRRKGIIFGHRDPE 61
QY 145 QVLDAVENKKPFYLYTGRGSEAMHVGHLIPFTTKMLQ-DVENVPLVIOMTDDEKYLW 203
DB 62 KILEKANGENVAVLTGMPGSK-FHGHKLTVQOLIYLOKNGFRV--FVALADEANAV 118
QY 204 KDLIDQAYGAVEN-AKDTACGFDIKT-FITSDLDYMGSSGFYKVVYKIQKHTFN 261
DB 119 RRIREEAVRIAYEYIANNMIALGIDPKDTEFFQ---TNRGPRYFLIOLFFSGKYAA 174
QY 262 QVKGIFG-FITSDICIGKISPAIOAAPSFSNSPQIFRDRDIOCLIPCAIDODPEYFMT 320
DB 175 EMEAILYELIYAKKMAASLT---QADILHVOLDYGGYR--HVVAVGVGADDDPHRLKT 227
QY 321 RDVAPR---IGYRPALHSTFFPALOGAOTKMSASDPNSIFLDTAKOIKTKVNKA 376
DB 228 RDLADRMAGVVELERPASTYHKLPGLDG--KKMSSSRDSTIFLDPREVAANKLPR-A 284
QY 377 FSGGRDITIEHRQFGNCDV-DVSEFMYITFLEDDDKLEQIRKDYTS---GAMLTGELKK 432
DB 285 LTGRATAEQRRLGCVPEVCVYHMDYHMLPDDGVEVKH---YTSCLRGKIIIGCECKQ 341
QY 433 ALIEVLOPLIAEHOARKEVTDIEYKEMFPR 464
DB 342 IAMEKLEFLEHOSRLEKAKTIAMKLVPR 373

RESULT 13
SYN_ARCFU STANDARD; PRT; 323 AA.
AC 029482;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tyrosyl-tRNA synthetase (EC 6.1.1.1) (Tyrosine-tRNA ligase) (TYRS).
OS TYRS OR AF0776.
GN Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;

RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin N., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utechtack T.,
RA Cotton M.D., Spriggs T., Artlach J.P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001051; AAB90462.1; -
DR TIGR: AF0776; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002307; tRNA-synt_tyr.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01040; TRNASYNTHTR.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 41 49 "HIGH" REGION.
FT SITE 214 218 "KMSKS" REGION.
SQ SEQUENCE 323 AA; 36616 MW; A655AEBA5116642 CRC64;

Query Match 7.5%; Score 192; DB 1; Length 323;
Best Local Similarity 23.1%; Pred. No. 3.9e-08;
Matches 78; Conservative 67; Mismatches 127; Indels 66; Gaps 15;

QY 139 SHRMNOVLDAVEKKKPFYLYTGRGSEAMHVGHLIPFTTKMLQVFNPIYQITDD 198
DB 19 TEEELRDLIEYKPR--AVGYEPGSE-THLGHMMTVOKLMDLQEA-GFEITVLADI 73
QY 199 EKYLMKDLTDQAVGAVENAKDILAGFDINKFTFISDDLDMGMSGFYKVVYKIQKHV 258
DB 74 HAYLNKGTFEELIENADVNNKVFIALGLDSRAKKFVLSGYQ-LSRDYVLDVAKMARIT 132
QY 259 TFCNVK---GIFGFTSDICIGKISPAIOAAPSFSNSPQIFRDRDIOCL-IPCA--- 310
DB 133 TLNRRARSMDVSRKREDPMVSCMITYPLMCA-----LDIAHLGVDLAVG 177
QY 311 IDQDPYFRMTDVAAPRIGYRPALHSTFFPALOGAOTKMSASDPNSIFLDTAKOIKT 370
DB 178 IDQKKIHMLARENLPRLGYSSPVLHPPIYGLDG--QKMSKSGN-YISRDPEEVER 234
QY 371 KVNK-HAFSG-----GRDTIEEHRQFGNCDVSEFMYITFLEDDDKLEQIRKDYTS 410
DB 235 KIRKAYCPAGVEENPLIDIAKTHILIRFGKIYVERAKRKG--DVE-----YASF----- 283
QY 411 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHOAR 448
DB 284 ---BELAEDFRSGQLHPLDKIAVAKYILNMLLEDARKR 318

RESULT 14
SYN_CIOLO STANDARD; PRT; 341 AA.
ID SYM_CIOLO
```



```

Db 113 RMSEFTKELALNEYITFVALGLDPEKINIVYLSQSKYKV-----KDLALISKRTMS 167
QY 262 QYKGIFGFTDSDICIGKISFPAIOAAPSFSNSFPQIFRDKT--DIOCLIFCALDODPYFMM 319
Db 168 EKKALIFGKGTETINIHVAPIVQVADIL--HPOLDENLSPPKRPVYVVGIDODPHRL 224
QY 320 TRDVAIPR---IGYPPALHSTFFPALOGAOTKMSASDPNSSIFLTDPAKOIKTKVKNKHA 376
Db 225 TRDIANKRAKEFFIPSPSYTHRFMTGLGG--KMSSSKRETAIFLTDDEKIVYKKILFS-A 281
QY 377 FSGGDTIEHRQFGG---NCDVVSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGELKKA 433
Db 282 KTGKRETLERHKYKGVPECVYVELFLY--HLILDKELAIETQKCRSGELTGCKKCM 339
QY 434 LIEVQPLAEHQARKREYTDIVK 458
Db 340 AYERVEEFLDKLKKERQAKETAVK 364

RESULT 9
SYM_METH STANDARD: PRT: 364 AA.
ID SYM_METH
AC 026352:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MTH251.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucelte-stramm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Potlier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA Mcounagall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000812; AAB84757.1; -
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_1b.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC PRINTS: PR01039; TRNASYNTTRP.
CC PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 70 "HIGH" REGION.
FT SITE 251 "KMSKS" REGION.
FT SITE 364 AA; 41301 MW; C2F348903338F61D CRC64;
SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

```

```

Query Match 15.6%; Score 397.5; Db 1; Length 364;
Best Local Similarity 27.9%; Pred No. 1,3e-24;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 85 VDPWTVTQSSANGIDYDKLIVRFSGSSKIDKELINRIERATGGRPHNIFLRGIFGSHRDN 144
Db 2 IDPW-----GSAG-LVEQDLINFGVRPF-SEVLDEV-----DEPSWLMRGIIIFGRDYE 50
QY 145 QYLDAYENKKRPYLYTGKRPSSAMHVGHLIFITFKVLQDYFNNPVLQYMTDDEKYLTK 204
Db 51 RIISAMKKEDEFAVYTGMPGSGR-MHIGHKMTVDQLRW-YDMGAEIIFPIDMEAYSAK 108
QY 205 DLTLOAVQDAVEN-AKDIACGFDINK-----TFIRSDIDYMGSSGFKNYVK 253
Db 109 GYDFEDSRRIAEIETIAGTIALGLDLEKDNINHYILOSEMLMEDLAY----- 156
QY 254 IOKHVTFNQVKGIFGFTSDICIGKISFPAIOAAPSFSNSFPQIFRDRDIOCLPCAIDQ 313
Db 157 LAGKVNENELRAIYGTGSTMAHMYAPILIQVSDILHPQDLDELGPR--PYIVVGPDO 213
QY 314 DPEFKTRDVAPRI-----GYPPALHSTFFPALOGAOTKMSASDPNSSIFLTDPAKOIK 369
Db 214 DPHIRLTRDIARFRDRYGFIPLSSYTHRFMTGLTG--KMSSNPKSAIFLSDPPEAE 271
QY 370 TVVKNHAFSGGDTIEHRQFGGNCVDVVSFMYLTFPLE-DDDKLEQIRKDYTSGAMLTG 428
Db 272 AKI-RNAKTGKRETLERHKGVECEIYETLLKHMGSQRLIEITYESCRNGLMG 330
QY 429 ELKALIEVQPLAEHQARKRE 451
Db 331 ECKNNTAEIRKFFELSVKREK 353

RESULT 10
SYM_ENCCU STANDARD: PRT: 134 AA.
ID SYM_ENCCU
AC 096771:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS) (Fragment).
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277683; PubMed=9615449;
RA Peyretallade E., Broussolle V., Peyret P., Metenier G., Gouy M.,
RA Vives C.P.;
RT "Microsporidia, amitochondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ012470; CAA10034.1; -
CC HSRP: P00952; ITC.
CC InterPro: IPR001412; tRNA-synt_1.
CC PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 89 "HIGH" REGION.

```

```
SYN_PYRAB
ID SYN_PYRAB STANDARD: PRT: 385 AA.
AC Q9UY11.
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR PAB111.
OS Pyrococcus abyssi.
OC Archaea: Euryarchaeota: Thermococcales; Thermococcaceae; Pyrococcus.
NCBI_TaxID=29292.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ248288; CAB50601.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Complete proteome.
KW SITE 82 "HIGH" REGION.
FT SITE 253 "KMSKS" REGION.
FT SEQUENCE 385 AA; 45100 MW; 4C29D01414976B12 CRC64;
SQ
Query Match 31.5%; Score 803; DB 1; Length 385;
Best Local Similarity 45.3%; Pred. No. 3,1e-57;
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;
OY 82 EDF-VDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRIERATQORPHHFLRGIFFSH 140
||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
3 EDFKVTPEVEGEV---VDYNKLIEHFGSPLETEELLEKTALTKSELPLFFRRKFFFSH 58
OY 141 RDMNOVLAYEKKRPFYLYTGRGSSSEAMVGHILPIFTKLDVFNVPVLYIOMTDEK 200
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
59 RQYDVLDVDEGRGFYLTGRGSPG-PMHIGHIIPFATKWLQKFGVNLTIQITLDEK 117
OY 201 YLMQD-LTLDAQVGDVANAQKDIAGCFDINKFTIFSDLYGMSSGCFYKNVAKIOKHT 259
: : : : : : : : : : : : : : : : : : : : : : : : : :
118 FLFKENLIFEDTKHWAYENIIDIIVGFDPOKTFIFQNSEF---TKIYEMAIPLAKKIN 173
OY 260 FNOVAGIGFTDSDCIGKISPPAIDAPSFSNPQIFRDRDTQICLIPCAIDODPYFRM 319
: : : : : : : : : : : : : : : : : : : : : : : : : :
174 FSMARVAGFTFEOSKIGMIFPAIDIAPTF-----FEKR---KCLIPALADODPYRNL 223
OY 320 TRDVPVIRGYRPAALHFFPALGAGTQKMSASDPNSSIFLTDPAKQIKRKYVNKHARG 379
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
224 QRDPAESLGYYKTAIHRSKVPSTLSGKMSASRPETALVITDSEDEKVMKFAALTG 283
OY 380 GRDTIEHRQEGCNDVVSFMYLTFLEDDDKBOIRKYD---TSGMALTGELKKALIE 436
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
284 GRPTLKQREKGEPEKCVFKEWLEIFFEEDK--KLKERYACAKNGELTGCECKRYLIS 341
OY 437 VLQPLIAHQARRKRYTDEIYK 458
```

```
Db 342 KIQEFLKEHOKRRKRAKQIEK 363
: : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 8
SYN_METJA
ID SYN_METJA STANDARD: PRT: 370 AA.
AC Q58810.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococcales; Methanococcaceae;
OC Methanococcus.
NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Cottonback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser K.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RT Science 273:1058-1073(1996).
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U67582; AAB99425.1; -
DR TIGR; MJ1415; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; FALSE NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Complete proteome.
KW SITE 75 "HIGH" REGION.
FT SITE 255 "KMSKS" REGION.
FT SEQUENCE 370 AA; 42660 MW; E6C71107CF82B59D CRC64;
SQ
Query Match 16.1%; Score 409.5; DB 1; Length 370;
Best Local Similarity 30.9%; Pred. No. 1.4e-25;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;
OY 87 PWTVOTSSAKGIDYDKLIVRFGSSKIDKELINRIERATQORPHHFLRGIFFSHDMNOV 146
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
8 PW-ETPAV--IDYKTKTEQEGVAKPIVDVLDQKE-----HFFRNIIILGHADFERI 57
OY 147 LDAYENKRPFYLYTGRGSSSEAMVGHILPIFTKWLQ---DVFNVPVLYIOMTDEKYL 203
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
58 VDAIKNNKEFNAVSGMMSGK-MHFGHKMAYVDLKFQYKTDNINIPF----ADLEAFAWA 112
OY 204 KDVLTDQAYGDV-ENAKDIAGCFDINKFTIFSDLYGMSSGCFYKNVAKI-QKHVTFN 261
```



```

QY 180 TKMLQDVENVPLVQMTEDEKYLKMDLTDQAYGDVAVENAKDIACGFDINKTFIFSDLD 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 TKMLQDVENVPLVQMSDEKYLKMDLTDQAYSTVENAKDIACGFDINKTFIFSDLE 243
QY 240 YMGSSGFYKVNKVIQKHVTEVNOVGIFGFTDSDCIGKISFPAIDAPFSNSPQIFRD 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 YMGSGPGFYKVNKVIQKHVTEVNOVGIFGFTDSDCIGKSSFPAYOAPSFSNSPKIFRD 303
QY 300 RTDIOCLIPCALIDDPYRPMRDVAPRIGYPRKALHSTFFPALOGACTKMSASPNSI 359
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 RTDIOCLIPCALIDDPYRPMRDVAPRIGYPRKALHSTFFPALOGACTKMSASPNSI 363
QY 360 FLUTYAKOIKTRVKNHAFSGGRDTEEHROFGNCDVDFSEMYLTFELEDDEKLEQIRKD 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 FLUTYAKOIKSVKNHAFSGGRDTEEHROFGNCEVDVSEMYLTFELEDDEKLEQIRKD 423
QY 420 YTSGAMLTGELKALIEVLQPLIAEHQARKREVDTEIYKEFTPRKLSFDFQ 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 424 YTSGAMLTGELKALIEVLQPLIAEHQARKREVDTEIYKEFTPRKLSFDFQ 475

```

RESULT 4

SYN_RABIT STANDARD: PRT: 475 AA.

```

ID 127 SYM_RABIT STANDARD: PRT: 475 AA.
AC P23612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TPRS).
GN MARS.
OS Euryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RX REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RP MEDLINE=94009008; PubMed=8404867;
RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
RA McCaughan K.R., Kisselev L.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins.";
RL EMO J. 12:4013-4019(1993).
CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1 SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -1 CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33460; AAA31246.1; ALT_SEQ.
CC EMBL: U02595; AAB60257.1; -.
CC PIR: A35904; YWRBPR.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synC_1b.
CC InterPro: IPR001412; tRNA-synC_1.

```

```

DR InterPro: IPR002306; tRNA-synC_trp.
DR Pfam: PF00579; tRNA-synC_1b; 1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00176; AA_TRNA_LIGASE_1; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68
FT SITE 168 177 WHEP-TRS.
FT SITE 353 357 "HIGH" REGION.
SO SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;

```

Query Match 86.1%; Score 2192; DB 1; Length 475;

Best Local Similarity 89.0%; Pred. No. 5,5e-169;

Matches 414; Conservative 24; Mismatches 27; Indels 0; Gaps 0;

```

QY 7 ASLELFNSIATOGELVSLKAGNASKDEIDSAVMYLSLMSYAAAGEDYKADCPGN 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 ASPQELFSSIAOGELVSLKARKAPKEIDSAVMYLSLKTYSKEAMGEDYKADCCPGN 70
QY 67 PAPTSNHGPDATAEADDPVDPWTVTSSAKGIDYDKLIVRFSSKIDELIRIRATGQ 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 STPDSHGDPDAVDKEDVDVDPWTVTSSAKGIDYDKLIVRFSSKIDELVIRATGQ 130
QY 127 RPHNPLRIGIFPSHRDMNOVDAYENKKPFYLYTGRGSSSEAHVGHILPFIETKMLQDV 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 RPHRFLRIGIFPSHRDMNOVDAYENKKPFYLYTGRGSSSEAHVGHILPFIETKMLQDV 190
QY 187 FNVPLVQMTEDEKYLKMDLTDQAYGDAVENAKITACGFDINKTFIFSDLDYMGSSG 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 FNVPLVQMSDEKYLKMDLTDQAYGYTLENNAKDIACGFDINKTFIFSDLDYMGSSG 250
QY 247 FKNVNVKIOKHVTEVNOVGIFGFTDSDCIGKISFPAIDAPFSNSPQIFRDRTDIOCL 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 FKNVNVKIOKHVTEVNOVGIFGFTDSDCIGKISFPAIDAPFSNSPQIFRGADIOCL 310
QY 307 IPCAIDODPYRPMRDVAPRIGYPRKALHSTFFPALOGACTKMSASPNSIPLTDYAK 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 IPCAIDODPYRPMRDVAPRIGYPRKALHSTFFPALOGACTKMSASPNSIPLTDYAK 370
QY 367 QIKTVKNHAFSGGRDTEEHROFGNCDVDFSEMYLTFELEDDEKLEQIRKDYSGAML 426
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 371 QIKTVKNHAFSGGRDTEEHROFGNCDVDFSEMYLTFELEDDEKLEQIRKDYSGAML 430
QY 427 TGELEKALIEVLQPLIAEHQARKREVDTEIYKEFTPRKLSFDFQ 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 431 TGELEKALIEVLQPLIAEHQARKREVDTEIYKEFTPRKLSFDFQ 475

```

RESULT 5

SYN_SCHPO STANDARD: PRT: 395 AA.

```

ID 009692;
AC 009692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TPRS).
GN SPAC2F.13C.
OC Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetes; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=972;
RA Gentile S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----

```

CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X53918; CAA37872.1; -
CC EMBL: X52113; CAA36356.1; -
CC EMBL: M74074; AAA30799.1; -
CC PIR: A40279; YWBO.
CC PIR: S14540; S14540.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC FT DOMAIN 24 69 WHEP-TRS.
CC FT SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
FT SITE 169 178 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT CONFLICT 17 17 L -> M (IN REF. 2).
SQ SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 89.8%; Score 2286.5; DB 1; Length 475;
Best Local Similarity 92.6%; Pred. No. 1.3e-176;
Matches 438; Conservative 12; Mismatches 20; Indels 3; Gaps 2;

QY 1 MNSERP--ASLELFNSITATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEY 58
DB 4 MSNGRGCGGSPLELFHSTIAAGELVRDLKARNAKDEIDSAVKMLSLKTSKATGEY 63
QY 59 KADCPGNPAPTSNHPDPATEAEEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 118
DB 64 KYDCPPGPAPESGEGDLATEDEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 123
QY 119 RIERTGGRPHRLRGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPFI 178
DB 124 RIERTGGRPHRLRGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPFI 183
QY 179 FTKMLQDVFNPLVQMTDDEKYLKDLTDQAYGDAVENAKDILACGFDIKTFIFSDL 238
DB 184 FTKMLQDVFNPLVQMTDDEKYLKDLTDQAYGDAVENAKD-ITCGFDINKTFIFSDL 242
QY 239 DYMGNSSGFYKNNVKIQKHVTENQVKGIFGFTSDSCICKISPPAIOAAPSFNSPQJIF 298
DB 243 DYMGNSSGFYKNNVKIQKHVTENQVKGIFGFTSDSCICKISPPAIOAAPSFNSPQJIF 302
QY 299 DRTDQCLIPALIDDPFRFRTRDVAPRIQYKPKALLHSTFPALQAGQTKKSADPNSS 358
DB 303 DRTDQCLIPALIDDPFRFRTRDVAPRIQYKPKALLHSTFPALQAGQTKKSADPNSS 362
QY 359 IFLDPTAQIKTKYKNAHAFSGGRDIEEHROFGNCDVDVSMYLTFFLEDDDKLEQJIR 418
DB 363 IFLDPTAQIKTKYKNAHAFSGGRDIEEHROFGNCDVDVSMYLTFFLEDDDKLEQJIR 422
QY 419 DYTSGAMLTGELKALIEVLOPLAIEHQARREYTDIVKEFMPRKLSFDFQ 471
DB 423 DYTSGAMLTGELKALIEVLOPLAIEHQARREYTDIVKEFMPRKLSFDFQ 475

AC P32921;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: Arg + L-tryptophan + tRNA(Trp) = AMP +
CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
CC ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
CC TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
CC EMBRYONIC STEM CELLS.
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X69656; CAA49347.1; -
CC EMBL: X69657; CAA49348.1; -
CC PIR: S31461; S31461.
CC PIR: S31462; S31462.
CC MGD: MGI:104630; Wars.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Alternative splicing.
CC FT DOMAIN 23 68 WHEP-TRS.
CC FT SITE 168 177 "HIGH" REGION.
CC FT SITE 353 357 "KMSKS" REGION.
CC FT VARSPLIC 476 481 MISSING (IN MAJOR ISOFORM).
SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;

Query Match 86.7%; Score 2208.5; DB 1; Length 481;
Best Local Similarity 89.0%; Pred. No. 2.6e-170;
Matches 420; Conservative 24; Mismatches 27; Indels 1; Gaps 1;

QY 1 MNSERP--ASLELFNSITATOGELVRSKAGNASKREIDSAVKMLVSLKMSYKAAAGEY 59
DB 4 MWSGESCTSPLELFNSITATOGELVRSKAGNAPKREIDSAVKMLVSLKMSYKAAAGEY 63
QY 60 ADCPPGNPAPTSNHPDPATEAEEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 119
DB 64 AGCPGNPTAGRNCDSDTKASEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 123
QY 120 IERATGGRPHRLRGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPFI 179
DB 124 IERATGGRPHRLRGIFFSHRDMQVLDAYENKKPFYLYTGRGSSSAMVGHILPFI 183

RA Frolova L.Y., Grigorieva A.Y., Sudomolina M.A., Kisseliev L.L.;
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-
 RT response elements and exon-intron organization.";
 RL Gene 128:237-245(1993).
 RN [7]
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.
 RC TISSUE=Keratinocytes;
 RC MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [8]
 RP FUNCTION.
 RA MEDLINE=92225128; PubMed=1373391;
 RA Bange F.-C., Flohr T., Buwilt U., Boeltger E.C.;
 RT "An Interferon-induced protein with release factor activity is a
 RT tryptophanyl-tRNA synthetase.";
 RL FEBS Lett. 300:162-166(1992).
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -I- SUBUNIT: HOMODIMER.
 CC -I- INDUCTION: BY INTERFERON GAMMA.
 CC -I- SIMILARITY: BELONGS TO CLASS-1 "WHEP-TRS" DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: M7804; AAA67324.1; -
 DR EMBL: X59892; CAA42545.1; -
 DR EMBL: M61715; AAA61298.1; -
 DR EMBL: X62570; CAA44450.1; -
 DR EMBL: S82905; AAB39381.1; -
 DR EMBL: X67920; CAB94198.1; -
 DR EMBL: X67921; CAB94198.1; JOINED.
 DR EMBL: X67922; CAB94198.1; JOINED.
 DR EMBL: X67923; CAB94199.1; -
 DR EMBL: X67924; CAB94199.1; JOINED.
 DR EMBL: X67925; CAB94199.1; JOINED.
 DR EMBL: X67926; CAB94199.1; JOINED.
 DR EMBL: X67927; CAB94199.1; JOINED.
 DR EMBL: X67928; CAB94199.1; JOINED.
 DR PIR: A41706; A41706.
 DR PIR: A41633; A41633.
 DR PIR: JH0533; JH0533.
 DR PIR: S19246; S19246.
 DR Aarhus/Ghent-2DPAGE; 3524; IEF.
 DR PHCI-2DPAGE; P23381; -
 DR MIM: 191050; -
 DR InterPro: IPR00738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b.1.
 DR Pfam: PF00458; WHEP-TRS.1.
 DR PRINTS: PRO1039; TRNASYNTTRP.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 KV Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 19 64 WHEP-TRS.
 FT SITE 164 173 "HIGH" REGION.
 FT SITE 349 353 "SKMSK" REGION.
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).
 FT CONFLICT 424 424 A -> R (IN REF. 4).
 SO SEQUENCE 471 AA: 53165 MW: E96344449053A0D0 CRC64;

Query Match 96.3%; Score 2454; DB 1; Length 471;
 Best Local Similarity 99.6%; Pred. No. 4,3e-190;
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MNSEPASLLELFNSTATGGLVRSIKAGNSKDEIDSAVKKLVSLKMSYKAAGEDYKA 60
 DB 1 MNSEPASLLELFNSTATGGLVRSIKAGNSKDEIDSAVKKLVSLKMSYKAAGEDYKA 60
 QY 61 DCPGPAPTSNHGPATAEEDFVDPTVOTSSAKDIDYDLTFRFSKIDKELINRI 120
 DB 61 DCPGPAPTSNHGPATAEEDFVDPTVOTSSAKDIDYDLTFRFSKIDKELINRI 120
 QY 121 ERATGGRPHFLRPGIFSFHRDMNOVLDAYENKKPFYLYTGGRPSSEAMHGHLPFT 180
 DB 121 ERATGGRPHFLRPGIFSFHRDMNOVLDAYENKKPFYLYTGGRPSSEAMHGHLPFT 180
 QY 181 KMLQDFNVPLVYQMTDDEKTYLMDLTDQAYGDAVENAKDILACGFDINKFTFSDLDY 240
 DB 181 KMLQDFNVPLVYQMTDDEKTYLMDLTDQAYGDAVENAKDILACGFDINKFTFSDLDY 240
 QY 241 MCMSSGFYKNVYKIOKHVFENOVKGIFGFTSDCIKISFPALIOAAPSNSFPQIFRDR 300
 DB 241 MCMSSGFYKNVYKIOKHVFENOVKGIFGFTSDCIKISFPALIOAAPSNSFPQIFRDR 300
 QY 301 TDIOCLIPCAIDODPYFRMTROVAPRIGYKRPALLHSTFPALOGAQTMSADPNSSIF 360
 DB 301 TDIOCLIPCAIDODPYFRMTROVAPRIGYKRPALLHSTFPALOGAQTMSADPNSSIF 360
 QY 361 LVDTAQQTITKYNKHAFFSGRPTIEHKGFGNCVDVDFMLTFFLEDDEKLEQIRXDY 420
 DB 361 LVDTAQQTITKYNKHAFFSGRPTIEHKGFGNCVDVDFMLTFFLEDDEKLEQIRXDY 420
 QY 421 TSGAMLTGELKALLLEVDLOPLAEHQARKKVTDEIVKFFMTPRKLSDFDQ 471
 DB 421 TSGAMLTGELKALLLEVDLOPLAEHQARKKVTDEIVKFFMTPRKLSDFDQ 471
 RESULT 2
 SYW_BOVIN STANDARD; PRT; 475 AA.
 ID SYW_BOVIN
 AC P17248;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrypRS).
 GN WARS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 ON NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Retina.
 RX MEDLINE=91329348; PubMed=1907847;
 RA Garret M., Pajot B., Trezequet V., Labouesse J., Merle M.,
 RA Gandar J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to
 RT prokaryotic synthetases but near identity with mammalian peptidic
 RT chain release factor.";
 RL Biochemistry 30:7809-7817(1991).
 RN [2]
 RP SEQUENCE OF 17-475 FROM N.A.
 RC TISSUE=Pancreas;
 RA Garret M., Trezequet V., Pajot B., Gandar J.-C., Merle M.,
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., la Bounessec B.,
 RA Labouesse J., Bonnet J.;
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -I- SUBUNIT: HOMODIMER.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 8.87395 Seconds
(without alignments)
2111.829 Million cell updates/sec

Title: US-09-813-718-10

Perfect score: 2547
Sequence: 1 MPNSEPASLELFENSIATQG.....KLSFDQKLAALHHNNHH 484

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2454	96.3	471	1	SYW_HUMAN
2	2286.5	89.8	475	1	SYW_BOVIN
3	2208.5	86.7	481	1	SYW_MOUSE
4	2192	86.1	475	1	SYW_RABIT
5	1210	47.5	395	1	SYW_SCHPO
6	1163	45.7	432	1	SYWC_YEAST
7	803	31.5	385	1	SYW_PYPAB
8	409.5	16.1	370	1	SYW_METJA
9	397.5	15.6	364	1	SYW_METTH
10	386	15.2	134	1	SYW_ENCCU
11	370.5	14.5	420	1	SYW_ARCFU
12	353	13.9	374	1	SYW_AERPE
13	192	7.5	323	1	SYW_ARCFU
14	178.5	7.0	341	1	SYW_CLOLO
15	172	6.8	394	1	SYWC_YEAST
16	170.5	6.7	366	1	SYW_SULSO
17	169	6.6	395	1	SYW_AQUAE
18	166.5	6.5	346	1	SYW_CHLTP
19	163.5	6.4	353	1	SYW_BORBU
20	160	6.3	346	1	SYW_CHLNU
21	159.5	6.3	337	1	SYW_TREPA
22	155	6.1	1714	1	SYEP_DROME
23	153	6.0	1440	1	SYEP_HUMAN
24	152	6.0	344	1	SYW_CHLTP
25	146	5.7	326	1	SYW_HELPY
26	138	5.4	319	1	SYW_METTH
27	137	5.4	528	1	SYW_HUMAN
28	136.5	5.4	528	1	SYW_BOVIN
29	135.5	5.3	328	1	SYW_THEMA
30	133	5.2	328	1	SYW_BACST
31	132.5	5.2	343	1	SYW_MYCLE
32	131	5.1	379	1	SYWM_YEAST
33	130	5.1	326	1	SYW_HELPY

34	129	5.1	401	1	SYWC_SCHPO	O14055 schizosacch
35	128	5.0	334	1	SYW_HAEN	P43835 haemophilus
36	124.5	4.9	347	1	SYW_MYCE	P47372 mycoplasma
37	122	4.8	350	1	SYWM_CAEPL	P46579 ctenorhadi
38	118.5	4.7	343	1	SYW_CLOAB	O97166 clostridium
39	116.5	4.6	330	1	SYW_BACSO	P21656 bacillus su
40	112.5	4.4	335	1	SYW_BUCAT	P57602 buchnera ap
41	111.5	4.4	336	1	SYW_MYCTU	O53386 mycobacteri
42	111.5	4.4	1067	1	IMB4_SCHPO	O60100 schizosacch
43	110	4.3	306	1	SYW_METJA	O57834 methanococc
44	107.5	4.2	1088	1	RRPO_ROTBR	P17468 bovine rota
45	107	4.2	330	1	SYW_RICPR	O92d76 rickettsia

ALIGNMENTS

RESULT 1	SYW_HUMAN	STANDARD:	PRT:	471 AA.
ID	P23381: P78535: Q9UDL3:			
AC	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--trna ligase)			
DE	(TrpRS) (IFP53) (hmrs).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92105071: PubMed=1761529;			
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92107982: PubMed=1763065;			
RA	Flecker J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9211058: PubMed=1765274;			
RA	Frlova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovleva O.L.,			
RA	Kisselev L.L.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92164636: PubMed=1537332;			
RA	Buitt U., Plohr T., Boettger E.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RN	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RX	MEDLINE=9631994: PubMed=8724762;			
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-			
RT	tRNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RN	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE=Spem;			
RX	MEDLINE=93292992: PubMed=7685728;			

Db 342 IAWKLERRFLAEHOSRLERAKTIAMKLVPPR 373

Search completed: October 24, 2002, 12:53:51
Job time : 18.2689 secs

Db 331 ECKNNTAEFIKKFEELSVMKREK 353

RESULT 13

Trypophan--tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)

C:Species: Encephalitozoon cuniculi

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43806

R:Payreallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P. Mol. Biol. Evol. 15, 683-689, 1998

A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene

A:Reference number: 422693; MUID:98277683

A:Accession: T43806

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PEV>

A:Cross-references: EMBL:AJ012470; PIDN:CA10034.1

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

C:Keywords: ligase

Query Match 15.2%; Score 386; DB 2; Length 134;

Best Local Similarity 51.5%; Pred. No. 4.8e-23;

Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 82 EDFDPMVTQTSAAK---GIDYDKLIYFGSSKIDKELINRIERATGQRPHHFLRGI 137

Db 3 EQRITPMQVEVYSDVEVAIDYDKIINQFCGEKFNQALADRLKLSKRAHYFRRGIV 62

QY 138 FSHRDMNOVDAYENKKPFYLYTGRPSSEAMVGHILPFIETKWLQDVNPVLVIOMTD 197

Db 63 FAHDFNLLDEIANNRPFLYTGGRPSKTMHIGHTIPFLCKYMQDAFKIRLIQTLD 122

QY 198 DEKYLWKDITLD 209

Db 123 DEKFLMKSRL 134

RESULT 14

E69461

tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C:Accession: E69461

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A:Authors: Uteerback, T.; Cotton, M.D.; Spriggs, T.; Artlisch, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon

A:Reference number: A69250; MUID:98049343

A:Accession: E69461

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 <KLE>

A:Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PIDN:AB89554.1; PID:9264885

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 14.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 4.2e-21;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 85 VDPMTVOTSSAKGIDYDKLIYFGSSKIDKELINRIERATGQRPHHFLRGI 144

Db 3 VTPMEVEGV---IDYSLIEFGMOPE-SEVLEPID---NPHILMRGAIGHRDYL 52

QY 145 QVLDAVENKKPFYLYTGRPSSEAMVGHILPFIETKWLQDVNPVLVIOMTDDEKYLWK 204

Db 53 RITAMOKKEKEMAVMVGFMPSG-LPHFGHKMTMDEIYHOSAGSKAFV-ATADMEASHVR 110

QY 205 DLTLDQAYGDAVENAKDIACGFDINKTFTFSLLDYMGSSGFYKNV-KIOKVTFTNOV 263

Db 111 GLSWEKTELMGLYIKSIIALGLREDAVIYFQS-----KSHVYKDLAFELSAEVNSEL 164

QY 264 KGIGFDTSDICIGKISFPALIAAPSFNSFPQIRDRDIOCLIPCAIDDPYRMRNDV 323

Db 165 RAIYGFNSDPSLAKMFYTAIOADTL---HPQLSDFGGPKRVVYPVAGADDPHMKLRDL 221

QY 324 APRI----- 327

Db 222 AARISIFSEFEVEGGVRYRSKGAELSLDLDEFDKKIEYEHNDIGEAEIERAVRKI 281

QY 328 -----GYRPPALLHSTFPALOGAQTMSASDPNSSIFLTPTAKQIKTVKNHATSG 380

Db 282 EVEIGCFAPFIPSSYTHRFITGLTG--KMSSSKPEYSISLLDPEEGAKKVMK-AFTGG 338

QY 381 RDTIEEHQFGCNDVDVSEFMYLTFPLED--DDKLEQIRKQDVTSGAMLTGELKKAILEYLO 439

Db 339 RATAEQRRLGCEPDRCVYFELYSFHLSDSEELNQLAECEBREGRLICGKKKMAELVK 398

QY 440 PLIAEHQARKEV 452

Db 399 SFLKEHQKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J. DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: D72477

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KAM>

A:Cross-references: DBJ:AP000064; NID:95105945; PIDN:BA81476.1; PID:95106165

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2461

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 13.9%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 8.2e-20;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 85 VDPMTVOTSSAKGIDYDKLIYFGSSKIDKELINRIERATGQRPHHFLRGI 144

Db 8 LDPV---GAVEIKDYDRLRTFCGIRPSEVL--PLRKAKMEPSPLMRGIIFGHRD 61

QY 145 QVLDAVENKKPFYLYTGRPSSEAMVGHILPFIETKWLQ-DVNPVLVIOMTDDEKYLW 203

Db 62 KILEAKARGERVAVALTFGFMPSGK-FHFGHKLTVDQLLYLQKNGKVV--FVAIADAEEFAV 118

QY 204 KDLTLDQAYGAVYNN-AKDIACGFDINKT-FITSDIDYMGSSGFYKNVKKIOKHTFN 261

Db 119 RIRIREEVRIAYEVIYANNMIALGLDPKDTETFIQ---TKRGIPYRLIQLSGSKYTAA 174

QY 262 QVKGIFG-FETSDSDICIGKISFPALIAAPSFNSFPQIRDRDIOCLIPCAIDDPYRMR 320

Db 175 EMEATVIGELTPAKKMASLT---QAQILNQLQDLEYGYR---HVVYPVAGADDPHRLRT 227

QY 321 RDVAPR---IGYRPPALLHSTFPALOGAQTMSASDPNSSIFLTPTAKQIKTVKNKA 376

Db 228 RDLADRMAGVYELRPRASTYHKLQPLGDD--RKMSSSRPSTIPLTPPEVARRKLR-A 284

QY 377 FSGGRDTIEEHQFGCNDV-DVSFMYLTFPLEDDDKLQIRKYTS---GAMLTGELK 432

Db 285 LTGGRATAEQRRLGCGVEVCSYVHMDLYHLMPPDDGEVKNH---YTCGRKLKIGCGCKQ 341

QY 433 ALIEVLOPLIAEHQARKEVTDLEIVKEPMR 464

Query Match 21.0% Score 535.5; DB 2; Length 380;
Best Local Similarity 35.2%; Pred. No. 4.6e-34;
Matches 142; Conservative 78; Mismatches 142; Indels 41; Gaps 17;

Oy AAEEDVDPMVOTSSAKIDYDKLIVRFSGSKIDKELINRIERATGQRPHFLRGIFFS 139
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db ADGNVPTTYANVESD---LDYEKLARARGADELDDQARRP-----DHPLVRNGLEYTA 53

Oy 140 HRDMNQVLDAVENKKPFYLVTGRGSEAMHVGHLIPFIETFKWLDVENVVLDTMTDE 199
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db GRDVDDFTLTAGEQS----IYTVGVSPGS--PMHLGHAMVFYFARRLDDEFGARVYVPLSDDE 108

Oy 200 KYLMKDILTDQAYGVGAVE--NAKDITACGFDPINKPIF---SDLDYM--GMSSGFYKNVVKI 254
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db KTFMFDQTPAET--GYLLNLANRLDLAVGFDEPELRIVVDTBDADVLYPLAFADV--- 164

Oy 255 QKHVFNQVKGIFGTGFTSDCGIKISFPALIOAPFSNSFPOLFRDRTOICLIPAIPOD 314
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db RHATLOWNYG-----EPDNWGQAFYPRAVOTAHLL---LPOLVHG--EHEFLVPLAVDOD 213

Oy 315 PYFRMTRDVAIRIGYP--KPALLHSTFPFALOQAOTKKMSASDPNSSIFLTJTAOKIKTKV 372
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db PHVRSRDIYAARKARPVPAGKPGALLMQFLPLSLAG--PGKMSSS--AGVISRLTDSPTVIREKV 271

Oy 373 NKHAFIGSRDITEEHROFGCMCDVDVSFWYLTFLIEDDD-KLEQIRKDYTSGAMLTGELK 431
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db RTHATGCRASVEEHRARGVPAEDVPFOYLSAFEPEPDAAETARIEREYRAGDDLSELK 331

Oy 432 KALLEVLOPLIAHQARKREVTDELIVEFMTRPKISPEFOKLA 474
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db DLAAIRTEIFAIAHORRRALGD--VTEALDAFAFRLTDEROORA 372

RESULT 11
F64476
tryptophan--tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii
N:Alternate names: tryptophanyl-tRNA synthetase
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: F64476
R:Built, C.J.: White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstein, K.G.; Merrick, J.M.; Glodek, A.;
rsion, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kalne, B.P.; Borodovsky, M.; Kleink, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999
A:Accession: F64476
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-370 <BU>
C:Genetics:
A:Map position: FOR1375865-1376997
A:Start codon: GTG
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog
C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 16.1%; Score 409.5; DB 2; Length 370;
Best Local Similarity 30.9%; Pred. No. 3.1e-24;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

Oy 87 PMVOTSSAKGIDYDKLIVRFSGSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOV 146
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 8 PW--ETPAV--IDYKTMEQGVKRIYDVLGDAKE-----HHFFRNIIIGHRDFERI 57

Oy 147 LDAYENKKPFYLVTGRGSEAMHVGHLIPFIETFKWLD--DVFNPLVIOMTDEKYLW 203
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db VDAIRNNKEFAVAVSAMPSGK-MHFGHKMVVDLKFQKYKTDININIP-----ADLEAWYA 112

Oy 204 KDILTDAQYGANV--ENADDITACGDDINKTFIFSOLDYMGSSGGYKANVAKI-QKHVTFN 261
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db RNMSSETTKELALNTINYIALGLDPERKINVYLOSCKOKV-----KDLAILSLKRTWS 167

OY	262	OYK:IGFTSDSDICKISFPALQAAPSFNSPFOIFDRDT--DIQCLIPCAIDDDPYFRM	319
		..: : :	..: : :
Db	168	EKMKAIFYGKGTENIGHVAFIYQVADIL--HPQDENLSEBPKKVYVPGVIDDDPHRL	224
		..: : :	..: : :
OY	320	TPDVAVR---IGYRPALLHSFPFALOGAOTKMSASPNISFLTDPAAKQIKKVNNA	376
		: :	: :
Db	225	TPDIANRRKKERKFIIPSSSTYHREMTGLLG--KMSSSKPEPAIFLTDEKIVKKIES-A	281
		: :	: :
OY	377	FSGGRDTEIEHRQFG---NCDVVSFMYLTFEEDDDKLPQIRKDYSGAMLGELKA	433
		: : : : :	: : : :
Db	282	KTGGEETIEEHRKYGQVEECVYELFLY--HLIIDKELAEIYQKCSGSLTGCKCKM	339
		: : :	: : :
OY	434	LIEVLQPLIAEHQARKEVTEIVK	458
		: : : : :	: : : :
Db	340	AERYVEFLDKERQAKETAVK	364
		: : : :	: : : :
RESULT 12			
E69131		tryptophan--tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain D	
N1		Alternate names: tryptophanyl-tRNA synthetase	
C:		Species: Methanobacterium thermoautotrophicum	
C:		Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999	
C:		Accession: E69131	
R:		Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T	
J:		Qiu, D.; Spadafora, R.; Vicaltre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,	
K:		S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.	
J:		Bacteriol. 179, 7135-7155, 1997	
A:		Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu	
A:		Reference number: A69000; MUID:98037514	
A:		Accession: E69131	
A:		Status: preliminary; nucleic acid sequence not shown; translation not shown	
A:		Molecule type: DNA	
A:		Residues: 1-364 <MTH>	
A:		Cross-references: GB:AE000812; GB:AE000666; NID:g2621296; PIDN:AB84757.1; PID:g262	
A:		Experimental source: Strain Delta H	
A:		Gene: MTH251	
C:		Genetics:	
A:		Start codon: TTG	
C:		Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo	
C:		Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis	
Query Match		15.6%; Score 397.5; DB 2; Length 364;	
Best Local Similarity		27.9%; Pred. No. 2.6e-23;	
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;			
OY	85	VDPMVQVSSAKGIDYDKLYRFGSSKIDKELINTEATGQRPNIHFRKGIFFSHRDN	144
		: :	: :
Db	2	IDPM---GSAG-LEYQDLIENTGVRPR-SEVLDEV-----PEPSWLMRRGIIIGHRYDE	50
		: :	: :
OY	145	OYLDVAYENKKRPVLTGTGGRPSEAMHVGHLIFGTITKLYDQVFNPLVIOGTTDDEKYLWK	204
		: :	: :
Db	51	RITSMKKGEEDVAVVTGMPISGR-MHIGHKMIVDQLR-VYMGAEITPLPADAEVASAR	108
		: :	: :
OY	205	DLTLQAYGDAVEN-AKDIACGPDINK-----TIFSLDDYMGSSGFEYKNVVK	253
		: :	: :
Db	109	CYDFPDSRRIAIEEYIAGYIALGDLDEKDNHIVYLOSENLAVEDLAY-----	156
		: :	: :
OY	254	IOKHVTEOVKQIFGFTSDSDICKISFPALQAAPSFNSPFOIFDRDTIOCLIPCAIDQ	313
		: :	: :
Db	157	LAKGVNFMELRIAYGFTGSTSMAMHYADIIQVSDILHQLDDELGPB---PVIVPVGPDQ	213
		: :	: :
OY	314	DYFPMTDVAPRI-----GYRPALLHSFPFALOGAOTKMSASPNISFLTDPAAKQIK	369
		: :	: :
Db	214	DPIHLTLDIARFDRORYGFIIPSSSTYHREMTGLTG--KMSSNPKPAISLDPPEAE	271
		: :	: :
OY	370	TPVNNHAFSGGRDTEIEHRQFGNCDDVDSFMYLTFLE--DDDKLEQIRKDYTSGAMLTG	428
		: :	: :
Db	272	AKI-RNAATGGGFTLEKQRELQGVPEECIIYETTLIHMSGSDSRLEIYESCRNGLMCG	330
		: :	: :
OY	429	ELKKALIEVLQPLIAEHQARKE	451
		: : : :	: : : :

Db 119 DDEKYMNEDEFTLDOTRSWAYNILDIIAIVGFPNPKTEFIDPDEYI---RNMYPITVAKIA 175
 Oy 256 KHVFNQKGIFFGTFDSDICIGKISFPALQAAFSNSFPQIRDRDFTDIOCLIPCAIDDDP 315
 Db 176 KKLATSEVRAFFGTFDASSNIGIFFPALQIAPT-----MEKK---RCLIPALDIDDP 225
 Oy 316 YFRMRDVAIRIGYKPKPALLHSTFFPALQGAQTCKMSASDPNSISILFTJAKOIKRVKMH 375
 Db 226 YWRLOORDIAESIGYKKAQIHKFLPPLTGREGKSSSNPEIATILVDDPKTVERKIKKY 285
 Oy 376 AFSGGRDTIEEHROFGNCDDVVSFWYLTFLEDDD-KLEQIRKDYTGAMLTGELKKAL 434
 Db 286 AFSGGQPTIELHRRYKGGNPEIDVPQWLYFFFEEDDNRIKIEIEEYRSRGMKLTGELKOIL 345
 Oy 435 IEVLQPLAEHQARKKEVTDDEIVKEFMTPRKLS 467
 Db 346 IDKLNLLEBHR-RRREAKELVHFXYDGKLA 377

RESULT 8

C75020
 tryptophanyl-tRNA synthetase (trps) PAB1111 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: C75020
 R:anonymous, Genoscope
 Submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A:Reference number: A75001
 A:Accession: C75020
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-385 <KAM>
 A:Cross-references: GB:A2248288; GB:AL096836; NID:g5458960; PIDN:CA50601.1; PID:g545911
 A:Experimental source: Strain Orsay
 C:Genetics:
 A:Gene: trps; PAB1111
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 31.5%; Score 803; DB 2; Length 385;
 Best Local Similarity 45.38; Pred. No. 5.7e-55;
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

Oy 82 EDF-DMPTVOTSSAKGIDYDKLIYRFGSSKIDKELINRERATGQRPHNLRGIFPSH 140
 Db 3 EDFKTPMHEGV---VDYKLIHFGTSPLEBELLEKTAELTSELPLFFRRKKFFFSH 58
 Oy 141 RDMNOVLDAENKKPFYLYTGRGSSSEAMHVGHLIPFIETKWLODVFNVLVQMTDDEK 200
 Db 59 RQYDKVLQDYEGRGFFLYTGRGSPG-PMHIGHIIPFATKWLOERKFGVNLXYQTDDCK 117
 Oy 201 YLMKD-LTLDOAYGAVENAKDIIACGFDINTFTFSOLDYMGMSGFGYKNVKKQKHVT 259
 Db 118 FLFKENLTFEDTKHAWYENILDIIVAGFDPDKTFTFQNSEF---TKIEEMALPKKIN 173
 Oy 260 FNOVVGIFGTFDSDICIGKISFPALQAAFSNSFPQIRDRDFTDIOCLIPCAIDDDPYFRM 319
 Db 174 FSKMAVAVGTFEGSKIGMIFFPALQIAPTF-----FEKK---RCLIPALDIDDPYRL 223
 Oy 320 TRDVAIRIGYKPKPALLHSTFFPALQGAQTCKMSASDPNSISILFTJAKOIKRVKHAFIG 379
 Db 224 QRFDAESIGYKKAQIHKFLPPLTGREGKSSSNPEIATILVDDPKTVERKIKKY 285
 Oy 380 GRDTIEEHROFGNCDDVVSFWYLTFLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 436
 Db 284 GRPLTKEQREKGEKCVFEMKLIFFFEEDDK-KLERYYACKNGELTGECCKRYLIS 341
 Oy 437 VLOPLAEHQARKKEVTDDEIVK 458
 Db 342 KIQEFLKEHQRRKKAQIEK 363

RESULT 9

G71206
 tryptophan--tRNA ligase (EC 6.1.1.2) - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C:Accession: G71206
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se
 M.; Ohkuni, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.; Ogu
 DNA Res. 5; 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
 A:Reference number: A71000; MUID:98344137
 A:Accession: G71206
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-301 <KAM>
 A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BA31046.1; PID:g3258363
 A:Experimental source: Strain OT3
 A:Note: this accession replaces an Interim accession for a sequence replaced by Genba
 C:Genetics:
 A:Gene: PH1921
 C:Superfamily: yeast tyrosine--tRNA ligase
 C:Keywords: aminocyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 24.68; Score 626.5; DB 2; Length 301;
 Best Local Similarity 44.28; Pred. No. 2.5e-41;
 Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

Oy 169 MHVGLIPFIETKWLODVFNVLVQMTDDEKLYMKD-LTLDOAYGDAVENAKDIIACGF 227
 Db 1 MHGIIHFFPATKWLOERFGVNLXYQTDDKFLKEKLTDDTKRMVYNDILDIIVAGF 60
 Oy 228 DINKTFISDDYMGMSGFGYKNVKKIQKHVFNQKGIFFGTFDSDICIGKISFPALQAA 287
 Db 61 DPKKTFIQNSEF-----TKIYEMALPKKINFSMAKAVFETEGSKIGMIFFPALQIAP 116
 Oy 288 SFSNSFPQIRDRDFTDIOCLIPCAIDDDPYFRMTRDVAIRIGYKPKPALLHSTFFPALQGAQ 347
 Db 117 TF-----FERK---RCLIPALDIDDPYWRLOORDIAESIGYKKAQIHKFLPPLTGREGKSSSNPEIATILVDDPKTVERKIKKY 166
 Oy 348 TKMSASDPNSISILFTJAKOIKRVKHAFIGSGRDTIEEHROFGNCDDVVSFWYLTFLE 407
 Db 167 GKMSASRPEIATILVDDPKTVERKIKKYGRPLTKEQREKGEKCVFEMKLIFF 226
 Oy 408 EDDDKLEQIRKDY---TSGAMLTGELKKALIEVLQPLAEHQARKKEVTDDEIVK 464
 Db 227 EEDDK-KLERYYACKNGELTGECCKRYLISKIQEFLKEHQRRKK-AEKLVEKFKYTG 283
 Oy 465 KLS 467
 Db 284 KLA 286

RESULT 10

G84373
 tryptophanyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G84373
 R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky
 ; Lelthuser, B.; Keller, K.; Cruz, R.; Danson, M.T.; Hough, D.W.; Maddocks, D.G.; Ja
 Jung, K.H.; Alam, M.; Freilas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483
 A:Accession: G84373
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-380 <STO>
 A:Cross-references: GB:AE004437; NID:g10581646; PIDN:AA620355.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: trps2
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 16.2689 Seconds
(without alignments)
2858.658 Million cell updates/sec

Title: US-09-813-718-10

Perfect score: 2547

Sequence: 1 MPNSEPALLELEFNSIATQG.....KLSPDFOKLAALHHHHHH 484

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 281338 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 281338

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:.*
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2454	96.3	471	1 A41706	tryptophan--trna 1
2	2286.5	89.8	475	1 YWBO	tryptophan--trna 1
3	2208.5	86.7	481	1 S50053	tryptophan--trna 1
4	2177	85.5	475	1 YWRBPR	tryptophan--trna 1
5	1210	47.5	395	2 S58157	hypothetical prote
6	1163	45.7	432	2 S51901	tryptophan--trna 1
7	907	35.6	386	2 C90190	tryptophan--trna 1
8	803	31.5	385	2 C75020	tryptophan--trna 1
9	626.5	24.6	301	2 G71206	tryptophan--trna 1
10	535.5	21.0	380	2 G84373	tryptophan--trna 1
11	409.5	16.1	370	2 E64476	tryptophan--trna 1
12	397.5	15.6	364	2 E69131	tryptophan--trna 1
13	386	15.2	134	2 T43806	tryptophan--trna 1
14	370.5	14.5	420	2 E69451	tryptophan--trna 1
15	353	13.9	374	2 D72477	probable tryptoph
16	274.5	10.8	513	2 F84371	tryptophan--trna 1
17	192	7.5	323	2 H69346	tyrosyl--trna synth
18	188	7.4	364	2 E72512	probable tyrosyl-t
19	185.5	7.3	341	2 D95260	tryptophan--trna 1
20	185.5	7.3	341	2 G98125	tryptophan--trna 1
21	174.5	6.9	341	2 H86633	tryptophan--trna 1
22	172	6.8	394	2 A45999	tyrosine--trna 11g
23	170.5	6.7	366	2 S75410	tyrosine--trna 11g
24	169	6.6	395	2 H70385	tryptophan--trna 1
25	168.5	6.6	895	2 A86410	protein F3M18.22 (
26	167	6.6	351	2 E75438	tryptophan--trna 1
27	166.5	6.5	346	2 B71496	tryptophan--trna 1
28	163.5	6.4	353	2 E70100	tryptophan--trna 1
29	163	6.4	460	2 C84750	probable tyrosyl-t

30	160	6.3	346	2 C81654	tryptophan--trna 1
31	159.5	6.3	337	2 F71300	tryptophan--trna 1
32	158.5	6.2	375	2 B75072	tyrosyl--trna synth
33	155	6.1	1714	1 S18644	multifunctional am
34	154.5	6.1	102	2 T44994	probable tryptoph
35	153	6.0	1440	1 SYH0GT	multifunctional am
36	152	6.0	344	2 H86590	tryptophan--trna 1
37	152	6.0	344	2 C72034	tryptophan--trna 1
38	149.5	5.9	327	2 C84374	tyrosyl--trna synth
39	149.5	5.9	408	2 T03741	probable tyrosine-
40	146	5.7	339	2 E64676	tryptophan--trna 1
41	146	5.7	337	2 F71093	tyrosine--trna 11g
42	145	5.7	337	2 A11066	tryptophan--trna 1
43	140.5	5.5	1149	2 T27567	hypothetical prote
44	138	5.4	319	2 H69102	tyrosine--trna 11g
45	135.5	5.3	328	2 C72370	tryptophan--trna 1

ALIGNMENTS

RESULT 1
A41706
tryptophan--trna 11gase (EC 6.1.1.2) [similarity] - human
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor h
C:Species: Homo sapiens (man)
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
C:Accession: A41633; A41706; S19246; JN0676; JH0533; S26287
R:Fleckner, J.; Rasmussen, H.H.; Justesen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9
A:Reference number: A41633; MUID:92107982
A:Accession: A41633
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:559892; NID:930820; PIDN:CAA42545.1; PID:930821
R:Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 24245-24248, 1991
A:Title: Interferon induces tryptophanyl-trna synthetase expression in human fibrobla
A:Reference number: A41706; MUID:92105071
A:Accession: A41706
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:M77804; NID:g184656; PIDN:AAA67324.1; PID:g184657
R:Buwilt, U.; Florin, T.; Boeltger, E.C.
EMBO J. 11, 489-496, 1992
A:Title: Molecular cloning and characterization of an interferon induced human cDNA w
A:Reference number: S19246; MUID:92164636
A:Accession: S19246
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423, 'R', 425-471 <BUW>
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:g32709
A:Note: 213-Ser and 214-Thr were also found
R:Frolova, L.Y.; Grigorjeva, A.Y.; Sudomoina, M.A.; Kisseliev, L.L.
Gene 120, 237-245, 1993
A:Title: The human gene encoding tryptophanyl-trna synthetase: Interferon-response el
A:Reference number: JN0676; MUID:93292992
A:Accession: JN0676
A:Molecule type: DNA
A:Residues: 1-141:182-471 <FROI>
A:Cross-references: GB:X67918; GB:S62837; NID:g37968; GB:X67919; NID:g37969; GB:X6792
4; NID:g37974; GB:X67925; GB:S62835; NID:g37975; GB:X67926; GB:S62856; NID:g37976; GB
A:Note: the authors translated the codon GGC for residue 55 as Cys and GAG for residu
R:Frolova, L.Y.; Sudomoina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kisseliev, L.L.
Gene 109, 291-296, 1991
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr
A:Reference number: JH0533; MUID:92112058
A:Accession: JH0533
A:Molecule type: mRNA

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SmithKline Beecham Corporation
;; STREET: 709 Swedeland Road
;; CITY: King of Prussia
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19406-0939
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/928,100
;; FILING DATE: 12-SEP-1997
;; CLASSIFICATION: 424
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 9619072.3
;; FILING DATE: 12-SEP-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gimmil, Edward R
;; REGISTRATION NUMBER: 38,891
;; REFERENCE/DOCKET NUMBER: P31624-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-270-4478
;; TELEFAX: 610-270-5090
;;
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 197 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-928-100-4
;;
Query Match 4.2%; Score 108; DB 3; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.0016;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;
;;
QY 307 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLSTFPALOGAOTKMSA 352
DB 5 VPVGTGDKPMIEQREIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG-NAKMSK 61
;;
QY 353 SDPNSSIFLDTAKQITKYNKHAFCGGRDIEHRQFGNCVDVSPMUTLTF--LEDD 410
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHIRVEDPGKIEGN---MVFHYLDVGFREDA 116
;;
QY 411 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDEI 456
DB 117 QEIDMKERYQRGIGDVKTKRYLLEIRLGPFRERRIEFKADM 162
;;
RESULT 13
US-09-183-134-2
; Sequence 2, Application US/09183134
; Patent No. 6165759
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/183,134
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/923,867
;; FILING DATE: 04-SEP-1997
;; APPLICATION NUMBER: 9619072.3
;; FILING DATE: 12-SEP-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gimmil, Edward R
;; REGISTRATION NUMBER: 38,891
;; REFERENCE/DOCKET NUMBER: P31624
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-270-4478
;; TELEFAX: 610-270-5090
;;
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 197 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-09-183-134-2
;;
Query Match 4.2%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.0016;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;
;;
QY 307 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLSTFPALOGAOTKMSA 352
DB 5 VPVGTGDKPMIEQREIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG-NAKMSK 61
;;
QY 353 SDPNSSIFLDTAKQITKYNKHAFCGGRDIEHRQFGNCVDVSPMUTLTF--LEDD 410
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHIRVEDPGKIEGN---MVFHYLDVGFREDA 116
;;
QY 411 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDEI 456
DB 117 QEIDMKERYQRGIGDVKTKRYLLEIRLGPFRERRIEFKADM 162
;;
RESULT 14
US-09-492-581-4
; Sequence 4, Application US/09492581
; Patent No. 6346409
; GENERAL INFORMATION:
; APPLICANT: Gentry, Danile
; APPLICANT: Greenwood, Claire
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6346409el trps
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/492,581
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,100
; FILING DATE: 12-SEP-1997
; APPLICATION NUMBER: 9619072.3

```

Db 153 S--RLSSGISFEFTYOILQSIDFY--TLHKHNIOLIGGADWGNITAGIDLIRKKE 207
QY 266 -----JRGFT-----DSDICIKISPPALQAPSPNSNPF--QIRDRDIDCLIPCAI 311
Db 208 GPRKAFVGLTIPLMLKADGKFGTAGATWLDPKITPEFYQFWINOD----- 258
QY 312 DDDPYFRMTDVAIRIGYPKAPALLHSTFFPALOGAQTMSASDPNSSIFLTDTAKOIKTK 371
Db 259 -----RDV---IKYLK-----FFTFLDKEE-----IDALAEKYEKE 286
QY 372 VKNHASSGGDTTIEHRQFGNCDVDVSEMYLTFELEDDEKLEQIKRDYTGAMLTGELK 431
Db 287 PKRR--EAQRRLAE-----VTRFVHDDALEEAOK--ISEALFSGNIK 326
QY 432 KALIEVLO-----PLIAHQARRKEVTDIEVKEFMTPRK 465
Db 327 DLITEIEGLEGHVPYV-ELTKDAKNIVDWLVDTIEPSK 365

```

```

RESULT 10
US-09-352-990-28
; Sequence 28, Application US/09352990
; Patent No. 6255090
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
; FILE REFERENCE: BB-1191
; CURRENT APPLICATION NUMBER: US/09/352, 990
; CURRENT FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092, 866
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 377
; TYPE: PRP
; ORGANISM: Synechocystis sp.
US-09-352-990-28

```

Query Match 4.3%; Score 109; DB 4; Length 377;
 Best Local Similarity 21.0%; Pred. No. 0.0037;
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;

```

QY 138 FSHRDMNQVLDAYENKKRFYLTGRGSSSEAMVGHILPFIPTKWLQDVENVLYIQMTD 197
Db 33 FAHRSTTAM-----DKPRIL-SGVOPGTGN-LHLGNLTGAI-RSWWEO-----QOHY 75
QY 198 DEKYLWKDL-----TLDOAYGDAVENAKDILACGFIDINKTF----- 235
Db 76 DNEFCVVDLAIYVPHNPQLAQ---DTLITIALYIACGIDLOXSTIFVQSHVAHSELA 132
QY 236 -----SDLDYMGSSGSEFYKNVVKIQRHVTFNQVKGITGFTSDSICIKISPPALQAPSF 289
Db 133 WLLNCTVPLMWLERMIOFKKAVKQGENVS-----VCLLDYPLVMAA--- 174
QY 290 SNEFQIFRRRTIOQL-----IPCALDQDYPFRMTRDV-----APRIGYKPA 333
Db 175 -----DILLYDADKVPREGQKHLELTRDVIIRINDKREGEDAPYLKLEPL 222
QY 334 L-LHSTFFPALOGAQTMSASDPN--SSIFLDTAKOIKTKVKNHAFSGGRDTEEHROF 390
Db 223 IRKEGARVMSLADGTTKMSKMSDESELSRINLDRPEMIKKKKVK----- 266
QY 391 GNCDDVDVSMYLTFFLEDDKLEQIRKDYTGAMLTGELKKAL----- 434
Db 267 ---CKTDPQ--RGLWF---DDPERECHNLLTYLLISNQTRKAVVADSCAEMGQFKPPL 319
QY 435 -----IEVLOPLIAHQARRKEVTDIEYK 458
Db 320 TETAIAALEPIQAKYAEITLADRGEL-DRIIO 349

```

```

RESULT 11
US-08-923-867-2
; Sequence 2, Application US/08923867
; Patent No. 5851809
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA SYNTHETASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923, 867
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: 9619072.3
; APPLICATION NUMBER: 12-SEP-1996
; FILING DATE: 12-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R.
; REGISTRATION NUMBER: 38, 891
; REFERENCE/DOCKET NUMBER: P31624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-867-2

```

Query Match 4.2%; Score 108; DB 2; Length 197;
 Best Local Similarity 25.3%; Pred. No. 0.0016;
 Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```

QY 307 IPCALDQDYPFRMTRD-----VAPRIGYKAPALLHSTFFPALOGAQTMSA 352
Db 5 VPGTQDKPMIEDQTRVRSFNNAAYNCVDLYVEPEGIYENE--RAGRPLQDQ-NKMKK 61
QY 353 SDPNSSIFLDTAKOIKTKVKNHAFSGGRDTEEHROFGNCDVDVSEMYLTFE--LEDQ 410
Db 62 S-LNNGIYLADADTLRKVKVSMYTPDRIYVDPCKIEGN---WVFHYLDVFGRPEDA 116
QY 411 DKLEQIRKDYTGAMLTGELKKALIEVLOPLIAHQARRKEVTDIE 456
Db 117 QETADMKERYQRGGLGDVKTGRYLLLETLERLQPIRRRIEFARKD 162

```

```

RESULT 12
US-08-928-100-4
; Sequence 4, Application US/08928100
; Patent No. 6046174
; GENERAL INFORMATION:
; APPLICANT: Gentry, Danile
; APPLICANT: Greenwood, Claire
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6046174el trps
; NUMBER OF SEQUENCES: 6

```


INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-743-130A-2

Query Match 6.6%; Score 167.5; DB 2; Length 409;
Best Local Similarity 20.6%; Pred. No. 1.8e-09;
Matches 79; Conservative 61; Mismatches 136; Indels 107; Gaps 16;

146 VLDAENK-KPELYTGRGSPSEAMHVGHLIPFI-----FTKWLQDV-----F 187
27 IKQVLEKENRPVKIYNGTATGPK-HGQYVPMKLAHFLKAGEVYVLLADLHAFIDNM 85
188 NPLVLIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKFTFISLDYMGSSGF 247
86 KADLEVYKRAKYEFVAKAIIKLSINPIERLKFVVGSSYQKGDDY---MDLFKLSNIV 142
248 YKNNVVK-----IKHVFNGVKGIFGFTSDDCIGKISEFPALQAPFSNSFPQIFRRTD 302
143 SONDARAGADVAKOVANPLISGLT-----YPLMOA-----IDEEHLG 180
303 IOCLIPCAIDODPYFMRTRDVAPRIGYKPKPALHSTFFPALQAGOTKMSADPNSSIFLT 362
181 VDAQFG-GVDQRIFFVLAENLPSIGYKRAHLMNPMVPGI-GGGGKMSADPNSSIDII 238
363 DTAQKQITKYNKNAFSGG--RDT-----IEHRQRCGN 393
239 EEPKRVKKNKYNASVACAGELKMDGLAFIEVYIPIAELKTGVEGAKLDIRPEKYG- 297
394 CDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMGTGELKALI---EVLOPLAEHQARR 449
298 ---PLST-----DSTEQLKADVVDGKLAPPLDKLGADKINLPIRAFEES- 343
450 KEVTDLVKEFEMTPRKLSPDFOR 472
344 -----EEFOVAQKNGYPEK 358

RESULT 6
US-09-357-251-37
Sequence 37; Application US/09357251
Patent No. 6271441
GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
APPLICANT: Famodu, Layo O.
APPLICANT: Orozco, Buddy
APPLICANT: Schwaber, James S.
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
FILE REFERENCE: BB-1193
CURRENT APPLICATION NUMBER: US/09/357,251
CURRENT FILING DATE: 1999-07-20
EARLIER APPLICATION NUMBER: 60/093,530
EARLIER FILING DATE: July 21, 1998
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO 37
LENGTH: 1440
TYPE: PRT
ORGANISM: Homo sapiens
US-09-357-251-37

Query Match 6.0%; Score 153; DB 4; Length 1440;
Best Local Similarity 51.7%; Pred. No. 5.9e-07;
Matches 30; Conservative 11; Mismatches 13; Indels 4; Gaps 1;

10 LELFNSIATGELVRSKAGNASKDEISAVKMLVSLKMSYKAAAGDYRADCPGPNP 67
679 LVILNRVAVGDDVRELKAKAPKEDVDAVAKQLLSLKAETKEKTGOEYK---PGNP 732

RESULT 7
US-08-705-868-4
Sequence 4; Application US/08705868
Patent No. 5885798
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,868
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0117 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1184699
US-08-705-868-4

Query Match 5.4%; Score 137; DB 2; Length 388;
Best Local Similarity 21.4%; Pred. No. 3.5e-06;
Matches 81; Conservative 64; Mismatches 151; Indels 83; Gaps 19;

130 HFLRGTI--FFSHRDMNOVDAYENKRPFLYTGRCSPSEAMHVGHLIPFI-FTKWLQDV 186
12 HLITRNGEVLGEKELKEILKERELK----IYNGTATGPK-HVAYVPMKSLADFLKAG 66
187 FNPPLVI---QMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKFTFISLDYMG 242
67 CEVYILFADLHAYLDNNKAPMELLEIVSYENVIKAM-DESIGVPLEKLTGTDYQ- 124
243 MSSGFYKNVYKIQKHVFNOVK---GIFGFTSDDCIGKISFPALQAPFSNSFPQIFR 298
125 LSKETVLDVYRLSSVVOHDSKAGAEVAKOVEHPILISGLLYPGLQ---LDEEYLV-- 179
299 DRDIDQCLIPCAIDODPYFMRTRDVAPRIGYKPKPALHSTFFPALQAGOTKMSADPNSS 358
180 ---DAQF---GQIDQRIFFFAEKYLPALGYSKRVHLMNPMVGLIS--SKMSSSEESK 231
359 IFLTDAKQITKYNKNAF-----SGRDTIEH-----RQFGNCVDV 397
232 IDLDRKEDVKKLKK-AFCEPGVNNNGVLSFTKHLVPLKSEFVILRDEKMGN---- 286
398 VSEMYLTFLEDDDKLEQIRKDYTSGLMGTGELKALIEVLOPLAEHQARRKEVDEIV 457
287 ---KTYTAYVD-----LEKDPRAEVVHPGDLKNSVEVALNKL-----DPIR 325

US-09-492-581-2

Query Match	7.3%;	Score 185.5;	DB 4;	Length 341;
Best Local Similarity	24.0%;	Pred. No. 1.4e-11;		
Matches	81;	Conservative	60;	Mismatches 128;
			Indels	69;
			Gaps	15

```

OY 154 KPELYYRGSSSAMVGHILPIEFKMYLODVENVVLVYOMIDEXK-LMKDITDIOAY 212
Dd 3 KPIIL-TGDRTYK-LHIGHYVSLKRR-----VLQEDKXMDYFVLEADQAL 49
OY 213 GDAVEN-----AKDIACGDFINTKFI--SDLDYMGMSGCFYNNV---KIQK 256
Dd 50 TDHAKDPOTIVESIGNVALDYLAGDIPDNSTFIQIOPELWELSYNNVLYSLARLER 1090
OY 257 HATFNQWKGJFGFTSDCIKISFPALQAA--PFSNSFPQIFRDRDTIOCLIPCAIDOD 3144
Dd 110 NPTVTELSQKGFESIPFTGFLVYRIADADITAFKANY-----VPVGTQK 1566
OY 315 PYFKMTD-----VAPRIGYKRPALHSHTEFFPALQOGAOTKMSADPNSSIF 360
Dd 157 PMIEOTREIYRSENNAYNCVULVEPEGIYRPN--RAGRLEPGIDG-NAKMSKS-LNNGIY 2122
OY 361 LTRPAKOIKTVKNNAHFSGSDTIEEHRQSGMCDVUVSEMYLTF--LEDDDKLEIRK 418
Dd 213 LADADLTFRKKVMSYMDRPHIRVEDPGKIEGN---MYNHIYDVFGRERDAQEIADMKE 268
OY 419 DYTSGAMITGELKALLEVDLOPLAEHQARKKEVTDEI 456
Dd 269 RYQGGGLADVTKRYULTEIERELGPIRRERIEPAKAM 306

```

RESULT 4
US-08-743-130A-39
; Sequence 39, Application US/08743130A

```

? GENERAL INFORMATION:
? APPLICANT: Sassanfar, Mandana
? APPLICANT: Gallant, Paul L.
? APPLICANT: Shen, Xiaoyu
? APPLICANT: Tao, Nianjun
? APPLICANT: Tao, Jianshi
? APPLICANT: Houman, Fariba
? TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE
? TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
? NUMBER OF SEQUENCES: 41
? CORRESPONDENCE ADDRESS:

```

```

COMPUTER READABLE FORM :
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A

```

ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CP195-12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS:

US-08-743-130A-39

Query Match	6.7%;	Score 169.5;	DB 2;	Length 409;
Best Local Similarity	20.6%;	Pred. No. 1.1e-09;		
Matches 79;	Conservative 61;	Mismatches 136;	Indels 107;	Gaps 16;

```

QY 146 VLDVYENK-KPEVLYTGRGPSESMHNGHLIPFI-----FTKMLQDV-----F 187
Db 27 IKDVLKENRPPVKIYTWGTAIPGRK-FCGYFPMYKILHFLKAGCEVYLLADLHAFIDNM 85
QY 188 NVPLVIMTDDKYLKMDLTLDQAYGDAVENAKDIIACGEDINKFTFESLDYWGMSGGF 247
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 KAPLEVYKRAKYEEFYVKALKISINPIERLKFVWGSSYQKGGDY---MDLFKISNIY 142
QY 248 YKNVVK-----IQHNVFYNQYKGIFFGTSDSCIGKISFPALQANPSSNSFPQIFRRTD 302
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 SONAKRAGADVYQVAVNPILSGI-----YPLMQA-----IDEHILG 180
QY 303 IQCLIPCALDDPFRMTRDVAPRIIGYKPKRLLHSTFFPALOGAQTMSADPNSIFLT 362
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 VDAQFG-GVDQRKFTVLAEEHLBPSIGIKKRAHLMNPVPGI-GGGKMSADPNSKIDII 238
QY 363 DTAQIKITKYVKNHAFSGG--RDT-----IEEHROGGN 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 EEPVYVKKVNSAVCARGELKNDLFIIEVYIQPIAEKLTGVEGAFKIDIREKXGG- 297
QY 394 CDVDSFMYLTFLELDDDKLEQIKRDYTSGLMTGELKALI---EVLQPLIAEHQARR 449
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 ---PLSY-----DSIEQLKADVDVGKLAPPDLSGVADKINELLAPIRAEFESS- 343
QY 450 KEVYDELVEKFMTPRKLSFDQK 472
Db 344 -----EEFQVAKNGKGYEYK 358

```

RESULT 5
US-08-743-130A-2
; Sequence 2, Application US/08743130A
; Patent No. 5871987

APPLICANT: Sassanfar, Mandana
 APPLICANT: Gallant, Paul L.
 APPLICANT: Shen, Xiaoyu
 APPLICANT: Tao, Mianjun
 APPLICANT: Tao, Jianshi
 APPLICANT: Houman, Fariba
 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:

```

:
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/743,130A
:

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CP195-12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540

[illegible]

RESULT 2
US-08-928-100-2
; Sequence 2, Application US/08928100

```

1      GENERAL INFORMATION:
2      APPLICANT: Gentry, Danlle
3      APPLICANT: Greenwood, Claire
4      APPLICANT: Lawlor, Elizabeth
5      TITLE OF INVENTION: No. 6046174e1 trps
6      NUMBER OF SEQUENCES: 6
7      CORRESPONDENCE ADDRESS:
8      ADDRESSEE: SmithKline Beecham Corporation
9      STREET: 709 Swedeland Road
10     CITY: King of Prussia
11     STATE: PA
12     COUNTRY: USA
13     ZIP: 19406-0939
14     COMPUTER READABLE FORM:
15     MEDIUM TYPE: Diskette
16     COMPUTER: IBM Compatible
17     OPERATING SYSTEM: DOS
18     SOFTWARE: FASTSEQ for Windows Version 2.0
19     CURRENT APPLICATION DATA:
20     APPLICATION NUMBER: US/08/928,100
21     FILING DATE: 12-SEP-1997
22     CLASSIFICATION: 424
23     PRIOR APPLICATION DATA:
24     APPLICATION NUMBER: 9619072.3
25     FILING DATE: 12-SEP-1996
26     ATTORNEY/AGENT INFORMATION:
27     NAME: Gimm!, Edward R
28     REGISTRATION NUMBER: 38,891
29     REFERENCE/DOCKET NUMBER: P31624-1
30     TELECOMMUNICATION INFORMATION:
31     TELEPHONE: 610-270-4478
32     TELEFAX: 610-270-5090
33     TELEX:
34     INFORMATION FOR SEQ ID NO: 2:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 341 amino acids
37     TYPE: amino acid
38     STRANDEDNESS: single
39     TOPOLOGY: linear
40     MOLECULE TYPE: protein
41     OS-08-928-100-2

```

	Matches	81, Conservative	60, Mismatches	128, Indels	69, Gaps	15,
QY	154	KPEVLYTGRGDSSEAMHVGHLPIETFKWLODFVNPVLYIOMTDDEKY-LMKDLTIDQAY	212			
Db	3	KPIIL-TGDPRTGK-LTHIGHYGSLKNR-----VLLQEEEDKXDMFVFLADQOL	49			
QY	213	GDAYN-----AKDIACGFDINKPTIF--SDLDYGMSSGFKNNV---KIQK	256			
Db	50	TDHADDPTQIVESIGNVALDYLAVALDLPKSTIFQSQIPELIAELSMYMNVLVSILARLR	109			
QY	257	HTVENQVKGIFGFTDSDCGIKSPFAIOAA--PSEFNSPQFLFRDRTDIQCLICAIPOD	314			
Db	110	NPTVKTETISQKGFGEISIPGFLVYPIAQAADITAKANT-----VPVGTDOK	156			
QY	315	PYFRMTRD-----VADRIQYPKDALLHSTFFPALQAGQTKMSASDPNSSIF	360			
Db	157	PMIEQTRFIVSEFNNAVNCODYLVEPEGIIPENE--RAGRLPGLDG-NAKMSKS-LNNGI	212			
QY	361	LTDFAKQIKTKVKNKAIFSGGRDTEIEHROFGGNCQDVDSFMULTFF--LEDDDKLEQIK	418			
Db	213	LADDLDTLRKKVMWYTPDPDHALRVDDPGKIEGN---WVFHYLDVFGREDQOEIADMK	268			
QY	419	DYTSGAMLTEGLKALIEVLOPLIAHOARREYVDEI	456			
Db	269	RYQRGCDGVKTKRYLLETLERELDPYRRRIEPAKDM	306			

RESULT 3
US-09-492-581-2

```

1  GENERAL INFORMATION:
2
3  APPLICANT: Gentry, Danile
4  APPLICANT: Greenwood, Claire
5  APPLICANT: Lawlor, Elizabeth
6  TITLE OF INVENTION: No. 6346409e1 trps
7  NUMBER OF SEQUENCES: 6
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: SmithKline Beecham Corporation
10 STREET: 709 Swedeland Road
11 CITY: King of Prussia
12 STATE: PA
13
14 COUNTRY: USA
15 ZIP: 19406-0939
16
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Diskette
19 COMPUTER: IBM Compatible
20 OPERATING SYSTEM: DOS
21 SOFTWARE: FASTSEQ for Windows Version 2.0
22
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/09/492,581
25 FILING DATE:
26
27 CLASSIFICATION:
28
29 PRIORITY APPLICATION DATA:
30 APPLICATION NUMBER: US/08/928,100
31 FILING DATE: 12-SEP-1997
32 APPLICATION NUMBER: 9619072.3
33 FILING DATE: 12-SEP-1996
34
35 ATTORNEY/AGENT INFORMATION:
36 NAME: Gimmil, Edward R
37 REGISTRATION NUMBER: 38,891
38 REFERENCE/DOCKET NUMBER: P31624-1
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE: 610-270-4478
41 TELEFAX: 610-270-5090
42
43 TELEX:
44
45 INFORMATION FOR SEQ ID NO: 2:
46
47 SEQUENCE CHARACTERISTICS:
48 LENGTH: 341 amino acids
49 TYPE: amino acid
50 STRADEDNESS: single
51 TOPOLOGY: linear
52
53 MOLECULE TYPE: protein
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 11.8319 Seconds
(Without alignments)
999.159 Million cell updates/sec

Title: US-09-813-718-10
Perfect score: 2547
Sequence: 1 MPNSEPASLELNFSTATQC.....KLSEDFOKLAALHHHHHH 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backlist1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	47.8	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	7.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	7.3	341	US-09-492-581-2	Sequence 2, Appl
4	169.5	6.7	409	US-08-743-130A-39	Sequence 39, Appl
5	167.5	6.6	409	US-08-743-130A-2	Sequence 2, Appl
6	153	6.0	1440	US-09-357-251-37	Sequence 37, Appl
7	137	5.4	388	US-08-705-868-4	Sequence 4, Appl
8	137	5.4	388	US-09-123-615-4	Sequence 4, Appl
9	113.5	4.5	418	US-08-855-910-11	Sequence 11, Appl
10	109	4.3	377	US-09-352-990-28	Sequence 28, Appl
11	108	4.2	197	US-08-923-867-2	Sequence 2, Appl
12	108	4.2	197	US-08-928-100-4	Sequence 4, Appl
13	108	4.2	197	US-09-183-134-2	Sequence 2, Appl
14	108	4.2	197	US-09-492-581-4	Sequence 4, Appl
15	104.5	4.1	370	US-08-415-593-45	Sequence 45, Appl
16	95	3.7	323	US-08-816-977-21	Sequence 21, Appl
17	95	3.7	418	US-08-844-054-2	Sequence 2, Appl
18	95	3.7	418	US-09-347-333-2	Sequence 2, Appl
19	94.5	3.7	877	US-08-907-166-8	Sequence 8, Appl
20	94	3.7	855	US-08-890-865A-10	Sequence 10, Appl
21	90.5	3.6	898	US-08-465-995A-4	Sequence 4, Appl
22	90.5	3.6	898	US-08-465-994C-4	Sequence 4, Appl
23	90.5	3.6	898	US-08-966-145-4	Sequence 4, Appl
24	90.5	3.6	920	US-08-101-593-4	Sequence 4, Appl
25	89.5	3.5	344	US-09-393-554-2	Sequence 2, Appl
26	89.5	3.5	898	US-08-465-995A-2	Sequence 2, Appl
27	89.5	3.5	898	US-08-465-994C-2	Sequence 2, Appl

28	89.5	3.5	898	2	US-08-966-145-2	Sequence 2, Appl
29	89.5	3.5	920	1	US-08-101-593-2	Sequence 2, Appl
30	89	3.5	719	1	US-08-082-849B-31	Sequence 31, Appl
31	89	3.5	719	5	PCT-US94-0162A-31	Sequence 31, Appl
32	89	3.5	2206	1	US-07-852-260-2	Sequence 2, Appl
33	89	3.5	2206	1	US-08-461-503-2	Sequence 2, Appl
34	89	3.5	2206	4	US-08-465-250-2	Sequence 2, Appl
35	88	3.5	606	2	US-08-883-534-3	Sequence 3, Appl
36	88	3.5	606	3	US-09-204-764-3	Sequence 3, Appl
37	87	3.4	428	3	US-08-331-625A-43	Sequence 43, Appl
38	87	3.4	1452	3	US-08-331-625A-2	Sequence 2, Appl
39	87	3.4	1452	5	PCT-US93-04384-18	Sequence 18, Appl
40	87	3.4	1452	5	PCT-US93-04692-2	Sequence 2, Appl
41	86.5	3.4	495	4	US-09-217-490-2	Sequence 2, Appl
42	86	3.4	1443	1	US-08-308-872B-2	Sequence 2, Appl
43	85	3.3	410	4	US-09-352-990-16	Sequence 16, Appl
44	83.5	3.3	510	1	US-08-249-112-3	Sequence 3, Appl
45	83.5	3.3	510	4	US-09-522-217-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1
US-08-876-885-26
Sequence 26, Application US/08876885
Patent No. 6174713
GENERAL INFORMATION:
APPLICANT: Shen, Xiaoyu
APPLICANT: Houman, Fatiba
TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-TRNA
TITLE OF INVENTION: SANTHEASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING
NUMBER OF INVENTIONS: SAME
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,885
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI97-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear
US-08-876-885-26

Query Match 47.8%; Score 1218.5; DB 4; Length 424;
Best Local Similarity 57.1%; Pred. No. 6.3e-124;
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

OY 78 TEAEDPEVPTVQ-----TSSAKGIDYKLVGRSSGKIDKELNRIERATGCRPHHFLR 133
DB 13 TESEDKITPWEVGAVVGRSGMDYDKLISQFGTKHITETLERFKQVYGEHPFLK 72

CC abyssal (see AAF86431 and AAF41223-7) and P. abyssal proteins. P. abyssal is
 CC a hyperthermophilic archaeon, which is isolated from deep-sea
 CC hydrothermal vents. The present sequence is one such P. abyssal protein.
 CC The proteins of the present invention have various potential industrial
 CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade.
 CC Note: This patent is in the same patent family as WO2000065062, which
 CC contains additional sequences as shown in AAB99132-AAB99143,
 CC AAF75903-AAH75920 and AAG66436.
 CC
 XX
 SQ Sequence 385 AA;
 Query Match 31.5%; Score 803; DB 22; Length 385;
 Best Local Similarity 45.3%; Pred. No. 2,2e-71;
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;
 QY 82 EDFVDPWVQTSANGIDYDKLIVRGSSKIDKELINRIERATGQRPNHFLRGJFESH 140
 DB 3 EDFKVPWVEVEGV---VDYKNKLIENHGTSPLTEELLEKTAELTKSELPLFFPRKFFESH 58
 QY 141 ROMNOYLDAYENKKPFYLYTGRGPSSAMHVGHLIPFTKWLQDVFNPLVLOMTDDEK 200
 DB 59 RQDKVLYQDYEGRGFLYTGRRPSC-PMHIGHLPFPATKWLQKFGVNLVQITDDEK 117
 QY 201 YLWKD-LTLDQAYDAVENAKDIACGFDINKTFIFSDLDYMGSSGFYKNVVKIOKHYT 259
 DB 118 FLKFNKTFEFDTHKMAVENLDIIAVGFDDPKTFIFONSEF---ITVYEMALPIAKIN 173
 QY 260 FNOVKIGFTSDCIKISFPAIQAAFSNSFPQIFRDRDIOCLIPCAIDDDPYFM 319
 DB 174 FSKAKVFGFTEOSKIGMIFFPALQIAPTF-----FEKR---RCLIPAIQDDPYWRL 223
 QY 320 TRQVAPRIGYKPFALLHSFEPALQGMOTKMSADPNSIFLIDTAKOITKYNKHAHFSG 379
 DB 224 QRPFAESLGYKTAHASKFVPSLTSLSGKMSASKPETALVLTDSPEVKKWKEALTG 283
 QY 380 GRDTIEHRQFGNCDVDVSMYLTFFLEDDDKLEQIRKDY---TSGAMLTGELKALIE 436
 DB 284 GRFLKEGRKGEPEKCVFKMLEIFFEEDK--KLKERYACKNGNELTCGCKRYLIS 341
 QY 437 VLQPLIAHQARKKEVTDIYK 458
 DB 342 KIOEFLKEHQRRKKAQKQIEK 363
 RESULT 15
 AAB58219
 ID AAB58219 standard; Protein; 142 AA.
 XX
 AC AAB58219;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 557.
 XX
 KM Human; Lung cancer associated protein; neuroprotective; cytosstatic;
 KM cardioactive; immunomodulatory; muscular active; vulnerrary;
 KM gastrointestinal; nephroproctric; antinfecitive; gynecological;
 KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KM proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200005180-A2.
 PD
 XX 21-SEP-2000.
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.

XX
 PI Ruben SM;
 XX
 DR WPI; 2000-587514/55.
 DR N-PSDB; AAF18095.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11; Page 1051; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548, lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytosstatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerrary; gastrointestinal
 CC general; nephroproctric; antinfecitive; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58349 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.
 XX
 SQ Sequence 142 AA;
 Query Match 22.8%; Score 579.5; DB 21; Length 142;
 Best Local Similarity 85.5%; Pred. No. 1.1e-49;
 Matches 118; Conservative 5; Mismatches 12; Indels 3; Gaps 2;
 QY 1 MPNSEPASLLEFNSTATOGELVRSKAGNASDEIDSAVKMLVSLKMSYKAAAGDYKA 60
 DB 5 MPNSEPASLLEFNSTATOGELVRSKAGNASDEIDSAVKMLVSLKMSYKAAAGDYKA 64
 QY 61 DCPGPNAPTSNHPDTEAEEDFVDPWVQTSANGIDYDKLIVRGSSKIDKELINRI 120
 DB 65 DCPGPNAPTSNHPDTEAEEDFVDPWVQTSANGIDYDKLIVRGSSXNNGELLXDX 124
 QY 121 ERATGQRPNHFLRGJF 138
 DB 125 E-STAKXTHS--GQGXFF 139

Search completed: October 24, 2002, 12:51:13
 Job time : 33.3193 secs

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142300.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

QY 194 QMTDDEKYLTKRDLTLDOAYGDAVENAKDIIACGFDINKFTIFSDLDYMGSSGFYKNVY 253
DB 151 QLTDEKSIWKMLSVESQSRRLARENAKDIIACGFDVTKTFTISDPDYVG--GAFYKNMWK 208
QY 254 IOKHVTNQVKKIFGFTSDSCIGKISFPALQAAAPSFNSFPQIFRDTYDIOCLIPCAIDQ 313
DB 209 VOKCVTLNKMAGIFGFSGEDPIAKLSFPVQAVSPSPSPHLPFGKLNKCLIPCAIDQ 268
QY 314 DPEYFMTDVAAPRGYPRPALHSTFPALQCAQTKMSASDPNSSIFLTDTAKOIKTKVN 373
DB 269 DPEYFMTDVAAPRLGYSKPALESTFPALQENKMSASDPNSAIVYTDASAKDIKNKIN 328
QY 374 KHAESGGRDTIEHRQFCNCDDVYSFMYLTFELEDKLEQIRKDYSGAMLTGELKKA 433
DB 329 RYAFSGGSDSIEKHELANEVLDPVXYLSFLEDDSELEHKEKEGEGRMLTGEVKKR 388
QY 434 LIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471
DB 389 LFEVLTEIVEKRRARRAVALTDEMDVAFMAVRPLPSKFE 426

RESULT 11

AAB66931

ID AAB66931 standard; Protein: 424 AA.

AC AAB66931;

DT 12-APR-2001 (first entry)

DE Tryptophanyl-tRNA synthetase.

KM Tryptophanyl-tRNA synthetase; enzyme.

OS Candida albicans.

PN US6174713-B1.

PD 16-JAN-2001.

PE 16-JUN-1997; 9705-0876885.

PR 16-JUN-1997; 9705-0876885.

PA (CUBI-) CUBIST PHARM INC.

PI Shen X, Houman F;

DR WPI; 2001-201806/20.

DR N-PSDB; AAF55855.

XX New nucleic acids encoding Candida cytoplasmic tryptophanyl-tRNA
PT (ctRNA) synthetases, useful for producing recombinant ctRNA synthetases
and detecting inhibitor of Candida ctRNA synthetase function -
XX Claim 4; Fig 1; 32pp; English.

XX The present sequence is Candida albicans cytoplasmic tryptophanyl-tRNA
CC (ctRNA) synthetase. ctRNA synthetase is useful for producing recombinant
CC ctRNA synthetases and detecting inhibitors of ctRNA synthetase.

SQ Sequence 424 AA;

Query Match 47.8%; Score 1218.5; DB 22; Length 424;

Best Local Similarity 57.1%; Pred. No. 6.8e-113;

Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 78 TEAEDEFVDPWTQ---TSSAKGIDYDKLIVFGSSKIDKELINRIERATGGRPHHFLR 133
DB 13 TESEEDKIFPWEVAGAVNDSKMGIDYDKLISQFKHITFEELERKQYTGGEPPHFLK 72
QY 134 RGIFFSHRDMNOVLAYENKRPVLYTGRGSSSEAMHVGHLIPFTKWLQDVFNPLVY 193
II:III III:III II: III:III III:III III:III III:III III:III

DB 73 RGVFSSQDLRLILDLYHGEPPFLYTGRRSSDSMHLGHMVPFIETKMLQEVFDPVLYI 132
QY 194 QMTDDEKYLTKR-DLTLDOAYGDAVENAKDIIACGFDINKFTIFSDLDYMGSSGFYKNVY 252
DB 133 ELTDEKFLFKHQLTIDDVKGFAENAMKDIIVAGFNPENTTIFSDLYMG--GAFYENVY 190
QY 253 KIOKHVTNQVKKIFGFTSDSCIGKISFPALQAAAPSFNSFPQIFRDTYDIOCLIPCAID 312
DB 191 RTSQITTTSTAKAAYGFGFSDSCIGKIHFASTQIATAFPSSFPDVLGLPPKTPCLIPCAID 250
QY 313 QDPYFMRDVAAPRGYPRPALHSTFPALQCAQTKMSASDPNSSIFLTDTAKOIKTKV 372
DB 251 QDPYFMRDVAADKRLFKPALHAKFPALQASTKMSASDPTTSTIFMGDTAKOIQKI 310
QY 373 NKHAESGGRDTIEHRQFCNCDDVYSFMYLTFELEDKLEQIRKDYSGAMLTGELK 432
DB 311 NKYAFSGGRATAEHRRELGNPEVDVAFQYLSFSDYDEKLAQLEQYRKGEILSGEMK 370
QY 433 ALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSF 468
DB 371 ECITVLOEFVSAYOERRSKYDDQYVERKMKPKLVF 406

RESULT 12

AAG23699

ID AAG23699 standard; Protein: 292 AA.

AC AAG23699;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 27102.

XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 9905-0121825.

PR 05-MAR-1999; 9905-0123180.

PR 09-MAR-1999; 9905-0123548.

PR 23-MAR-1999; 9905-0125788.

PR 25-MAR-1999; 9905-0126264.

PR 29-MAR-1999; 9905-0126785.

PR 01-APR-1999; 9905-0127462.

PR 06-APR-1999; 9905-0128234.

PR 08-APR-1999; 9905-0128714.

PR 16-APR-1999; 9905-0129845.

PR 19-APR-1999; 9905-0130077.

PR 21-APR-1999; 9905-0130449.

PR 23-APR-1999; 9905-0130510.

PR 28-APR-1999; 9905-0130891.

PR 30-APR-1999; 9905-0131449.

PR 30-APR-1999; 9905-0132048.

PR 30-APR-1999; 9905-0132407.

PR 04-MAY-1999; 9905-0132484.

PR 05-MAY-1999; 9905-0132485.

PR 06-MAY-1999; 9905-0132486.

PR 06-MAY-1999; 9905-0132487.

PR 07-MAY-1999; 9905-0132863.

PR 11-MAY-1999; 9905-0134256.

PR 14-MAY-1999; 9905-0134218.

PR 14-MAY-1999; 9905-0134219.

PR 14-MAY-1999; 9905-0134221.

PR 14-MAY-1999; 9905-0134370.

PR 18-MAY-1999; 9905-0134768.

PR	18-JUN-1999	9905-01394.55
PR	18-JUN-1999	9905-01394.55
PR	18-JUN-1999	9905-01394.55
PR	18-JUN-1999	9905-01394.57
PR	18-JUN-1999	9905-01394.58
PR	18-JUN-1999	9905-01394.60
PR	18-JUN-1999	9905-01394.60
PR	18-JUN-1999	9905-01394.61
PR	18-JUN-1999	9905-01394.62
PR	18-JUN-1999	9905-01394.63
PR	18-JUN-1999	9905-01394.63
PR	18-JUN-1999	9905-01397.50
PR	21-JUN-1999	9905-01397.63
PR	21-JUN-1999	9905-01398.17
PR	22-JUN-1999	9905-01398.99
PR	23-JUN-1999	9905-01403.93
PR	23-JUN-1999	9905-01403.54
PR	24-JUN-1999	9905-01406.95
PR	28-JUN-1999	9905-01408.23
PR	29-JUN-1999	9905-01409.81
PR	30-JUN-1999	9905-01412.87
PR	01-JUL-1999	9905-01414.22
PR	01-JUL-1999	9905-01414.254
PR	02-JUL-1999	9905-01420.55
PR	06-JUL-1999	9905-01423.90
PR	08-JUL-1999	9905-01428.03
PR	09-JUL-1999	9905-01428.20
PR	12-JUL-1999	9905-01429.77
PR	13-JUL-1999	9905-01433.524
PR	14-JUL-1999	9905-01435.424
PR	15-JUL-1999	9905-01440.05
PR	16-JUL-1999	9905-01440.85
PR	16-JUL-1999	9905-01440.86
PR	19-JUL-1999	9905-01443.33
PR	19-JUL-1999	9905-01443.32
PR	19-JUL-1999	9905-01443.35
PR	19-JUL-1999	9905-01443.352
PR	20-JUL-1999	9905-01446.84
PR	21-JUL-1999	9905-01448.14
PR	21-JUL-1999	9905-01450.86
PR	21-JUL-1999	9905-01450.88
PR	22-JUL-1999	9905-01450.87
PR	22-JUL-1999	9905-01450.875
PR	22-JUL-1999	9905-01450.89
PR	22-JUL-1999	9905-01451.92
PR	23-JUL-1999	9905-01451.43
PR	23-JUL-1999	9905-01451.45
PR	23-JUL-1999	9905-01452.18
PR	23-JUL-1999	9905-01452.24
PR	26-JUL-1999	9905-01455.76
PR	27-JUL-1999	9905-01455.18
PR	27-JUL-1999	9905-01455.18
PR	27-JUL-1999	9905-01455.18
PR	28-JUL-1999	9905-01455.91
PR	02-AUG-1999	9905-01463.86
PR	02-AUG-1999	9905-01463.88
PR	02-AUG-1999	9905-01465.89
PR	03-AUG-1999	9905-01470.38
PR	04-AUG-1999	9905-01472.04
PR	04-AUG-1999	9905-01471.92
PR	05-AUG-1999	9905-01473.60
PR	06-AUG-1999	9905-01473.60
PR	06-AUG-1999	9905-01474.16
PR	09-AUG-1999	9905-01474.93
PR	09-AUG-1999	9905-01475.35
PR	10-AUG-1999	9905-01481.71
PR	11-AUG-1999	9905-01481.71
PR	12-AUG-1999	9905-01483.41
PR	13-AUG-1999	9905-01485.65
PR	13-AUG-1999	9905-01486.84

PR	15- ANG-1999	9905-0149368
PR	17- ANG-1999	9905-0149175
PR	18- ANG-1999	9905-0149426
PR	20- ANG-1999	9905-0149722
PR	20- ANG-1999	9905-0149723
PR	20- ANG-1999	9905-0149929
PR	20- ANG-1999	9905-0149902
PR	23- ANG-1999	9905-0149930
PR	23- ANG-1999	9905-0149930
PR	25- ANG-1999	9905-0150566
PR	26- ANG-1999	9905-0150584
PR	27- ANG-1999	9905-0151065
PR	27- ANG-1999	9905-0151066
PR	27- ANG-1999	9905-0151080
PR	30- ANG-1999	9905-0151103
PR	31- ANG-1999	9905-0151348
PR	01- SEP-1999	9905-0151363
PR	07- SEP-1999	9905-0152360
PR	10- SEP-1999	9905-0153708
PR	13- SEP-1999	9905-0154018
PR	15- SEP-1999	9905-0154739
PR	16- SEP-1999	9905-0154739
PR	20- SEP-1999	9905-0154739
PR	22- SEP-1999	9905-0155139
PR	23- SEP-1999	9905-0155486
PR	24- SEP-1999	9905-0155639
PR	28- SEP-1999	9905-0156458
PR	29- SEP-1999	9905-0156596
PR	04- OCT-1999	9905-0157117
PR	05- OCT-1999	9905-0157753
PR	06- OCT-1999	9905-0157865
PR	07- OCT-1999	9905-0158029
PR	08- OCT-1999	9905-0158232
PR	12- OCT-1999	9905-0158369
PR	13- OCT-1999	9905-0159283
PR	13- OCT-1999	9905-0159294
PR	13- OCT-1999	9905-0159295
PR	14- OCT-1999	9905-0159329
PR	14- OCT-1999	9905-0159330
PR	14- OCT-1999	9905-0159330
PR	14- OCT-1999	9905-0159637
PR	14- OCT-1999	9905-0159638
PR	18- OCT-1999	9905-0159584
PR	21- OCT-1999	9905-0160741
PR	21- OCT-1999	9905-0160768
PR	21- OCT-1999	9905-0160768
PR	21- OCT-1999	9905-0160770
PR	21- OCT-1999	9905-0160814
PR	21- OCT-1999	9905-0160815
PR	22- OCT-1999	9905-0160980
PR	22- OCT-1999	9905-0160981
PR	22- OCT-1999	9905-0160989
PR	25- OCT-1999	9905-0161404
PR	25- OCT-1999	9905-0161405
PR	25- OCT-1999	9905-0161406
PR	26- OCT-1999	9905-0161359
PR	26- OCT-1999	9905-0161360
PR	26- OCT-1999	9905-0161361
PR	28- OCT-1999	9905-0161920
PR	28- OCT-1999	9905-0161920
PR	28- OCT-1999	9905-0161943
PR	29- OCT-1999	9905-0162142

	Query Match	51.2%	Score 1304	DB 21	Length 426
	Best Local Similarity	60.3%	Pred. No. 1.9e-121		
	Matches 240	Conservative 67	Mismatches 87	Indels 4	Gaps 2
Q7	76 DATEAE--EDFDPMTVYQSSAKGIDYKLLIRFGSSKIDKELINRIEATGQORPHNEIR	133			
	: : :				
D8	31 DEREAESSQYVNPWEVSASKDGKIDYDKLIDKFGQRIDESLIDRVORTLSRQPHVFLR	90			
Q7	134 RGIFSRHSDMNOVLDAYLEAKRFYLITGSGPSSAMNHVHLIRPFTKKLQDVFNPLVI	193			
	: :				
D8	91 RSVFPAHRRFNDILDAEYSGDKFYLITGSGPSSALNHLHLIRPMFTKKLQDAFVPLVI	150			

```
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.
```

Query Match 51.2%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 1,7e-121;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

```
OY 76 DATEAE--EDFVDPWTVOTSSAKKIDYDKLIVFGSSKIDKELIRIERATGQRPHHFLR 133  
DB 7 DEREAESSQVAVNPWEVSADKDGKIDYDKLIDKFCORLDESLIDRVOLTRSRQPHVFLR 66  
OY 134 RGIFPSHRDMNOVLDAYENKRFELYTGKPSSEAMHGHILPFIFETWLODVENPLVI 193  
DB 67 RSVFPAHRDPNFIIDAYEKGDFLYITGRGSEADHGLHLPFETWLYLODAEFVPLVI 126  
OY 194 QMTDEKYLKMDLTDQAYGAVENAKDIIACGFDINTFTIFSDDYMGSSGFYKNVVK 253  
DB 127 QLTDEKSIWKMLSVESQRLARENAKDIACGFDVTKTFIFSDPDYVG--GAFYKNVVK 184  
OY 254 IOKHTFNOVKIRFGFTSDCGKISFPAIQAPSFNSFQIFEDRDIDQICLCAIDQ 313  
DB 185 VKGCVTLNKAMKIFGFSGDPIAKLSFPVQAVPSFSSFPHLFPGKDNRLCLICAIQ 244  
OY 314 DPEFMTDVAARIIGYPRFALHSTFPALOGAQTMSADPNSISFILTDAKQITKTVN 373  
DB 245 DPEFMTDVAARLGYSKPALESTFPALOGENKMSADPNSAIYVTDKAKDINKNIN 304  
OY 374 KHAESGSDGTTEHRQSGNCDVVSFMYLTFELEDQIRKDYTSGAMLNGELKA 433  
DB 305 RYAFSGQDSIEKIRHRELANLEVIDIPVKYLSFLEBDESELEHIRKEYGEGRLTGEVKKR 364
```

```
OY 434 LIEVLQPLIAEHQARRKEVTEIVEKFMTPRKLSDFQ 471  
DB 365 LIEVLTEIVEKRRARAAVTDEMVAFAVRLPSKFE 402  
RESULT 10  
AAG23697  
ID AAG23697 standard; Protein: 426 AA.  
XX AAG23697;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27100.  
XX  
KW Protein identification: signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX  
PM EP1033405-A2.  
XX  
ED 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123160.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128274.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139492.
```

KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143642.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

OY	371	KUNNAHSGEDTTEEHROFGNCDVDSMUJLFELEDDCKLEQIRKDYTSAMLTGEL	430
Dd	302	KUNNAHSGGDTEEHROFGNCDVDVSUMLJFLELEDDCKLEQIRKYDTSAMLTGEL	361
OY	431	KKALIEVLORPILAHNOARKKEUYDELUKEMTRJKLSDFDKLAALENNINNN	484
Dd	362	KKALIEVLORPILAHNOARKREUYDELUKEMTRJKLSDFDKLAALENNINNN	415

RESULT 6	
AAB47618	
ID	AAB47618 standard; Protein; 392 AA..

AC AAB47618;

DT 07-JAN-2002 (first entry)

DE Human inactive TRPRS.

KM Tyrosyl-RNA synthetase; TYRS; Rossmann fold nucleotide binding domain;
KM vascular endothelial cell function; burn; plastic surgery; abdomen;
KM polymorphonuclear leucocyte elastase; angio genesis; tumour metastasis;
KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KM dermal ulcer; diabetic ulcer; endothelialization;
KM tryptophanyl-tRNA synthetase; trPRS; vascular graft surgery.

OS Homo sapiens.

PN WO200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US089666.

PR 31-MAR-2000; 2000US-193471P.

PA (Scripps Res Inst.

PI Schimmel P, Wakasugi K;

DR WPI; 2001-626377/72.

DR N-PSDB; AAH43605.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
XX
XX
XX Disclosure; Page 135-36; 150pp; English.
XS

The sequences given in ABA47615-18 show full length and truncated versions of triptophanyl-tRNA synthetase (TrpRS). The truncated TrpRSs and the invention comprises a Rossmann fold nucleotide binding domain, and is capable of regulating vascular endothelial cell function. It is of approx. 40 Kilo Dalton molecular weight and is produced by cleavage of full length TrpRS with polymorphonuclear leucocyte elastase. Truncated TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing angiogenesis to a graft, treating myocardial infarction, solid tumor, and a condition that would benefit from increased or decreased angiogenesis in a mammal, in particular humans. It is also useful in diagnosis and as a wound healing agent for treating wounds such as dermal ulcers, diabetic ulcers, burns and injuries and in plastic surgery when reconstruction is required following a burn or for cosmetic purposes. It is particularly useful in the treatment of abdominal wounds where there is high risk of infection. Truncated TrpRS promotes endothelialization in vascular graft surgery and is used in conjunction with angiography to administer the angiogenic tRNA synthetase polypeptides or polynucleotides directly to the lumen and wall of the blood vessel.

SQ Sequence 392 AA;

Query Match	81.1%;	Score 2066;	DB 22;	Length 392;
Best Local Similarity	100.0%;	Pred. No. 9.1e-198;		

```
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	94	SAKIDYDKLIVFEGSSKIDKELINTEATORHNHRLRGIFSFHDMNQVDAEAK	153
Db	2	SAKIDYDKLIVFEGSSKIDKELINTEGRATORHNHRLRGIFSFHDMNQVDAEAK	61
Qy	154	KRFYLTGRGSPSEAHVGHILRFETKVLQDVFNVPVLIOMTDEKXMKDILDOAYG	213
Db	62	KRFYLTGRGSPSEAHVGHILRFETKVLQDVFNVPVLIOMTDEKXMKDILDOAYG	121
Qy	214	DAYNADITACGFIDINKITFISDLDYMGMSGFKNNVKLOKIVTENQYKGI6GFIDSD	273
Db	122	DAYNADITACGFIDINKITFISDLDYMGMSGFKNNVKLOKIVTENQYKGI6GFIDSD	181
Qy	274	CIGKISFPALQAPSPNSNPROIFRRJRDIOCLICALIDORUYRMYRDAVRI6YKPA	333
Db	182	CIGKISFPALQAPSPNSNPROIFRRJRDIOCLICALIDORUYRMYRDAVRI6YKPA	241
Qy	334	LHSTFPALQAGOTMSASDRPNSSIFLTDYKQIKITLVKNNHAFSGGRDITEEHROGSG	393
Db	242	LHSTFPALQAGOTMSASDRPNSSIFLTDYKQIKITLVKNNHAFSGGRDITEEHROGSG	301
Qy	394	CDVUVSFMULTFPLEDDDKLEOIRKDYSGAMLTGELKALIEVLORPLAHOARKEUT	453
Db	302	CDVUVSFMULTFPLEDDDKLEOIRKDYSGAMLTGELKALIEVLORPLAHOARKEUT	361
Qy	454	DELVEKFMTRKLSRDFOKLAALENNHNH	484
Db	362	DELVEKFMTRKLSRDFOKLAALENNHNH	392

RESULT 7
ABB64621
ID ABB64621 standard; Protein; 430 AA

AC ABB64621;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 20655.

KW *Drosophila*; developmental biology; cell signalling; insecticide;

XX

XX
XX

XXXXXXXXXXXXXXX

[illegible]

PD 27-SEP-2001
XX

XX PE ZS-MAK-ZUOL; ZUOLWC-US09Z231.

PR 11-JUN-2000: 2000HS-0614150
PR 23-MAR-2000: 2000US-19163/P.

XX
XX
(DEFE) DE CORP NY

XX
DT

11/25/2001

XX
2
EET
0001
CCCCC
7E

DR N-PSDB; ABL08724.

PT New isolated nucle

PT interactions -

PS Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention

XX The sequences given in AAB47615-18 show full length and truncated
 CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
 CC the invention comprises a Rossmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kDa molecular weight and is produced by cleavage of
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic tRNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.

XX
 SQ Sequence 437 AA;

Query Match 91.2%; Score 2324; DB 22; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-223;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

48 MSYAAAGEDYKACDPPNPAPTSNHPDPAEEDPDPKVTOTSSAKGIDYDKLYRF 107
 1 MSYAAAGEDYKACDPPNPAPTSNHPDPAEEDPDPKVTOTSSAKGIDYDKLYRF 60
 108 GSSKIDKELINRERATGQRPHHRLRGIFFSHRDMNOVLDAVENKKPFYLYTGRPSSE 167
 61 GSSKIDKELINRERATGQRPHHRLRGIFFSHRDMNOVLDAVENKKPFYLYTGRPSSE 120
 168 AMYGHILPFTFKLQDVENVPLVITQMTDDEKYLKRLTLDOAYGDAVENAKDIIACGF 227
 121 AMYGHILPFTFKLQDVENVPLVITQMTDDEKYLKRLTLDOAYGDAVENAKDIIACGF 180
 228 DINKTFISDDYMGMSGCFKKNVYKIQKHTFNQVKGIFGFTSDSCIGKISPAIQANP 287
 181 DINKTFISDDYMGMSGCFKKNVYKIQKHTFNQVKGIFGFTSDSCIGKISPAIQANP 240
 288 SFSNSFPQIFRDRDIOCLICAIIDQDYPFRMTDVAIRIGYPRPALHSTFFPALQGAQ 347
 241 SFSNSFPQIFRDRDIOCLICAIIDQDYPFRMTDVAIRIGYPRPALHSTFFPALQGAQ 300
 348 TKMSASDPNSSIFLDTAKQIKTKVKNKHAIFSGGRDTIEHHQFGNCDVDSFMTLTFEL 407
 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAIFSGGRDTIEHHQFGNCDVDSFMTLTFEL 360
 408 EDDDKLEQIRKDYTSAGMLTGEKALKALIEVLOPLAEHQARKREYTDIVKEFMTPRKLS 467
 361 EDDDKLEQIRKDYTSAGMLTGEKALKALIEVLOPLAEHQARKREYTDIVKEFMTPRKLS 420
 468 FDFOKLAALAEHHNNH 484
 421 FDFOKLAALAEHHNNH 437

RESULT 5
 AAB47617
 ID AAB47617 standard; Protein: 415 AA.
 XX
 AC AAB47617:
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human supermini TrpRS.
 XX
 KM Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;

KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KM dermal ulcer; diabetic ulcer; endothelialization;
 KW triphophanyl-tRNA synthetase; trpRS; vascular graft surgery.
 XX
 OS Homo sapiens.
 XX
 PN WO200174841-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 21-MAR-2001; 2001WO-US08966.
 XX
 PR 31-MAR-2000; 2000US-193471P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PI Schimmel P, Wakasugi K;
 XX
 DR WPI; 2001-626377/72.
 DR N-PSDB; AAA43604.
 XX
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction
 XX
 PS Disclosure; Page 129-30; 150pp; English.

XX
 CC The sequences given in AAB47615-18 show full length and truncated
 CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
 CC the invention comprises a Rossmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kDa molecular weight and is produced by cleavage of
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic tRNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.

XX
 SQ Sequence 415 AA;

Query Match 86.1%; Score 2194; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.5e-210;
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

71 SNHGPDAEAEEDVPDVTOTSSAKGIDYDKLIVRGSSKIDKELINRERATGQRPHH 130
 2 SNHGPDAEAEEDVPDVTOTSSAKGIDYDKLIVRGSSKIDKELINRERATGQRPHH 61
 131 FLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHGHILPFTFKLQDVENVNP 190
 62 FLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHGHILPFTFKLQDVENVNP 121
 191 LVITQMTDEKYLKRLTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGMSGCFYKN 250
 122 LVITQMTDEKYLKRLTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGMSGCFYKN 181
 251 VVKIQKHTFNQVKGIFGFTSDSCIGKISPAIQAAASFSNSFPQIFRDRDIOCLIPCA 310
 182 VVKIQKHTFNQVKGIFGFTSDSCIGKISPAIQAAASFSNSFPQIFRDRDIOCLIPCA 241
 311 IDDPYFRMTDVAIRIGYPRPALHSTFFPALQGAOTKMSASDPNSSIFLDTAKQIKT 370
 242 IDDPYFRMTDVAIRIGYPRPALHSTFFPALQGAOTKMSASDPNSSIFLDTAKQIKT 301

QY 181 KWLQDVNPLVLTQMTDEKYLKDLTLDQAYGDAVENAKDIACGFDINKTFISDLDY 240
 DB 185 KWLQDVNPLVLTQMTDEKYLKDLTLDQAYGDAVENAKDIACGFDINKTFISDLDY 244
 QY 241 MGSSSGYKKNVVKIQKHVTENQVKGIFGFTSDICIGKISFPALQAAFSNSFPQIFRDR 300
 DB 245 MGSSSGYKKNVVKIQKHVTENQVKGIFGFTSDICIGKISFPALQAAFSNSFPQIFRDR 304
 QY 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIF 360
 DB 305 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIF 364
 QY 361 LTTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDY 420
 DB 365 LTTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDY 424
 QY 421 TSGAMLTGELKALIEVLQPLIAEHQARKEVDEIVKEFMTPRKLSFDRQ 471
 DB 425 TSGAMLTGELKALIEVLQPLIAEHQARKEVDEIVKEFMTPRKLSFDRQ 475

RESULT 3

AA05372
 ID AAY05372 standard; Protein; 471 AA.

AC AAY05372;

DT 30-JUN-1999 (first entry)

DE Human HCMV inducible gene protein, SEQ ID NO 12.

KW HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;
 KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;
 XX drug screening.

OS Homo sapiens.

PN WO9913075-A2.

PD 18-MAR-1999.

PF 08-SEP-1998; 98WO-US18638.

PR 22-SEP-1997; 97OS-0059725.

PR 08-SEP-1997; 97US-0058180.

PA (UYPR-) UNITV PRINCETON.

PI Cong J, Schenk T, Zhu H;

DR WPI: 1999-243729/20.

DR N-PSDB; AAX33942.

PT New isolated human genes

PS Claim 3; Page 112-114; 184pp; English.

XX This sequence is encoded by a human gene of the invention, and is induced
 CC to express by both HCMV and interferon (IFN), designated HCMV-inducible
 CC genes (c1g or c1gs). The invention also relates to genes that are
 CC repressed in the presence of HCMV infection, designated HCMV-repressible
 CC genes (crg or crgs). The products can be used to obtain agents which can
 CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
 CC also be used for the development of drugs that would allow for higher
 CC dosage IFN treatments without the concomitant toxicity normally
 CC associated with administering high levels of IFN. The products can also
 CC be used for detection, diagnosis and drug screening.

XX Sequence 471 AA;

Query Match 96.2%; Score 2449; DB 20; Length 471;
 Best Local Similarity 99.4%; Pred. No. 5.7e-236;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MPNSEPASLLEFNSIATOGELVSLKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60
 DB 1 MPNSEPASLLEFNSIATOGELVSLKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60
 QY 61 DCPGPNAPPTSNHGPDPATEAEEDFVDPWVOTSSAKGIDYDKLIVRGSSKIDKELINRI 120
 DB 61 DCPGPNAPPTSNHGPDPATEAEEDFVDPWVOTSSAKGIDYDKLIVRGSSKIDKELINRI 120
 QY 121 ERATGQRPHFLRGRGFEFSHRDMNOVDAYENKRPFLYGRGSSPAMHGHILPFT 180
 DB 121 ERATGQRPHFLRGRGFEFSHRDMNOVDAYENKRPFLYGRGSSPAMHGHILPFT 180
 QY 181 KWLQDVNPLVLTQMTDEKYLKDLTLDQAYGDAVENAKDIACGFDINKTFISDLDY 240
 DB 181 KWLQDVNPLVLTQMTDEKYLKDLTLDQAYGDAVENAKDIACGFDINKTFISDLDY 240
 QY 241 MGSSSGYKKNVVKIQKHVTENQVKGIFGFTSDICIGKISFPALQAAFSNSFPQIFRDR 300
 DB 241 MGSSSGYKKNVVKIQKHVTENQVKGIFGFTSDICIGKISFPALQAAFSNSFPQIFRDR 300
 QY 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIF 360
 DB 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIF 360
 QY 361 LTTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDY 420
 DB 361 LTTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDY 420
 QY 421 TSGAMLTGELKALIEVLQPLIAEHQARKEVDEIVKEFMTPRKLSFDRQ 471
 DB 421 TSGAMLTGELKALIEVLQPLIAEHQARKEVDEIVKEFMTPRKLSFDRQ 471

RESULT 4

AAB47616
 ID AAB47616 standard; Protein; 437 AA.

AC AAB47616;

DT 07-JAN-2002 (first entry)

DE Human mini T1PRs.

KW Tyrosyl-tRNA synthetase; TYRS; Rossmann fold nucleotide binding domain;
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
 KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KM dermal ulcer; diabetic ulcer; endothelialization;
 KM tryptophanyl-tRNA synthetase; trprs; vascular graft surgery.

XX Homo sapiens.

PN WO200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US08966.

PR 31-MAR-2000; 2000US-193471P.

PA (SCRI) SCRIPPS RES INST.

PI Schimmel P, Makasugi K;

DR WPI: 2001-626377/72.

DR N-PSDB; AAH43603.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

PT vascular endothelial function, in particular for regulating

PT angiogenesis, tumor metastasis and treating myocardial infarction -

PS Disclosure; Page 123-24; 150pp; English.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
XX angiogenesis, tumor metastasis and treating myocardial infarction
XX
PS Disclosure; Page 117-19; 150pp; English.
XX
CC The sequences given in AAB47615-18 show full length and truncated
CC versions of triphosphanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.

XX
XX
SQ Sequence 484 AA;

Query Match 100.0%; Score 2547; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 9.2e-246;

Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPNSEPASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60
DB 1 MPNSEPASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60
QY 61 DCPGPNAPRTSNHGPDATAEEDFVDPWTVOYSSAGIDYDKLYVFGSSKIDKEILNRI 120
DB 61 DCPGPNAPRTSNHGPDATAEEDFVDPWTVOYSSAGIDYDKLYVFGSSKIDKEILNRI 120
QY 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLPIFT 180
DB 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLPIFT 180
QY 181 KWLQDFVNLVLTOMTDEDEYLYMKDLTLDOAYGDAVENANDIACGFDINKTEIFSDLDY 240
DB 181 KWLQDFVNLVLTOMTDEDEYLYMKDLTLDOAYGDAVENANDIACGFDINKTEIFSDLDY 240
QY 241 MGSMSGFYKVVNIQKHVFNQYKGFEGFDSQIGKISPPALQAAFSFNSFPQIFRDR 300
DB 241 MGSMSGFYKVVNIQKHVFNQYKGFEGFDSQIGKISPPALQAAFSFNSFPQIFRDR 300
QY 301 TDIOCLIPCAIDDDPYFRMTDVAAPRIGYKPPALHSTFPALQGAQTAKSADPNSSIF 360
DB 301 TDIOCLIPCAIDDDPYFRMTDVAAPRIGYKPPALHSTFPALQGAQTAKSADPNSSIF 360
QY 361 LDTAKOIKTKVKNHAFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 420
DB 361 LDTAKOIKTKVKNHAFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 420
QY 421 TSCAMLTGELKKALILEVLOPLIAEHQARRKEVDETVKEMTPRKLSFPOKLAALLENH 480
DB 421 TSCAMLTGELKKALILEVLOPLIAEHQARRKEVDETVKEMTPRKLSFPOKLAALLENH 480
QY 481 HHHH 484
DB 481 HHHH 484

RESULT 2
AAB58220
ID AAB58220 standard; Protein; 475 AA.
XX

AC AAB58220;
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polypeptide sequence SEQ ID 558.
XX

KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnery;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN WO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.

PI Ruben SM;

DR WPI: 2000-587514/55.

DR N-PSDB; AAF18096.

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.

PS Claim 11; Page 1052-1053; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX associated proteins and polynucleotide sequences, their agonists, and
XX antagonists may have neuroprotective; cytostatic; cardioactive;
XX immunomodulatory; muscular active general; vulnery; gastrointestinal
XX general; nephrotropic; antiinfective; gynecological; or antibacterial
XX activity. The invention also includes antibodies specific for the
XX protein or polynucleotide sequences. The lung cancer associated
XX polynucleotide sequences may be used for detection of lung cancer,
XX chromosome identification, as chromosome markers, and for numerous other
XX diagnostic or research purposes. The proteins may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders. The proteins may also be used in the treatment of wounds and
XX infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
XX peptide AAB58549 are used in the course of the invention for the
XX identification and characterisation of the polynucleotide and protein
XX sequences.

SQ Sequence 475 AA;

Query Match 96.3%; Score 2454; DB 21; Length 475;
Best Local Similarity 99.6%; Pred. No. 1.8e-236;

Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPNSEPASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60
DB 5 MPNSEPASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 64
QY 61 DCPGPNAPRTSNHGPDATAEEDFVDPWTVOYSSAGIDYDKLYVFGSSKIDKEILNRI 120
DB 61 DCPGPNAPRTSNHGPDATAEEDFVDPWTVOYSSAGIDYDKLYVFGSSKIDKEILNRI 124
QY 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLPIFT 180
DB 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLPIFT 184

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 30.3193 seconds
(without alignments)
1773.117 Million cell updates/sec

Title: US-09-813-718-10

Perfect score: 2547
Sequence: 1 MPNEPASILFELFNSTATQG.....KLSPDFOKLAALHHHHHH 484

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_032802:*

1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2547	100.0	484	22	AAB47615 Human full-length
2	2454	96.3	475	21	AAB58220 Lung cancer associ
3	2449	96.2	471	20	AAV05372 Human HCMV Inducib
4	2324	91.2	437	22	AAB47616 Human mini TyRS.
5	2194	86.1	415	22	AAB47617 Human supermini tr
6	2066	81.1	392	22	AAB47618 Human inactive Trp
7	1378	54.1	430	22	ABBB64621 Drosophila melanog
8	1378	54.1	430	22	ABBB67203 Arabidopsis thalia
9	1304	51.2	402	21	AAG23698 Arabidopsis thalia
10	1304	51.2	426	21	AAG23697 Arabidopsis thalia
11	1218.5	47.8	424	22	AAB66931 Typtophanyl-trna

12	968	38.0	292	21	AAG23699 Arabidopsis thalia
13	831	32.6	173	21	AAB58517 Lung cancer associ
14	803	31.5	385	22	AAB96409 Putative P. abyss
15	579.5	22.8	142	21	AAB58219 Lung cancer associ
16	269	10.6	71	22	AAO10441 Human polyptide
17	185.5	7.3	341	19	AAW56423 Typtophanyl trna
18	185.5	7.3	341	22	AAU37851 Streptococcus pneu
19	173	6.8	372	22	AAB47614 Human mini TyRS m
20	171.5	6.7	372	22	AAB47612 Human mini TyRS.
21	169.5	6.7	409	20	AAW94248 C. albicans tyrosy
22	167.5	6.6	409	20	AAW94247 C. albicans tyrosy
23	159.5	6.3	379	22	AAV37623 Protein involved i
24	158.5	6.2	378	20	AAV96600 Putative P. abyss
25	155	6.1	1714	22	ABBB1487 Drosophila melanog
26	153	6.0	1512	22	AAW78732 Human protein SEQ
27	153	6.0	1512	22	AAW79716 Mammalian toxicolo
28	153	6.0	1550	22	AAV35439 Human protein SEQ
29	152	6.0	344	20	AAV35439 Chlamydia pneumoni
30	146	5.7	339	22	AAU35829 Helicobacter pylor
31	144	5.7	525	22	ABBB60745 Arabidopsis thalia
32	138.5	5.4	140	21	AAAG23032 Arabidopsis thalia
33	137	5.4	536	22	AAB47611 Human full length
34	128	5.0	334	22	AAU35477 Haemophilus influe
35	123	4.8	347	21	AAAG31904 Arabidopsis thalia
36	122.5	4.8	346	21	AAAG48617 Arabidopsis thalia
37	120.5	4.7	404	22	AAU54020 Propionibacterium
38	115.5	4.5	129	21	AAAG23033 Arabidopsis thalia
39	113.5	4.5	418	22	AAE02578 Enterococcus tyros
40	113.5	4.5	423	22	AAU35085 Enterococcus faeca
41	113	4.4	415	22	AAU35998 Staphylococcus aur
42	113	4.4	420	18	AAW19781 Tyrosyl-trna synth
43	113	4.4	420	22	AAU37056 Staphylococcus aur
44	112	4.4	345	22	AAAG90498 C glutamicum prote
45	108	4.2	334	22	AAU38213 Salmonella typhl c

ALIGNMENTS

RESULT 1	AAAB47615	standard; Protein; 484 AA.
ID	AAAB47615	standard; Protein; 484 AA.
XX	AAAB47615:	
AC	07-JAN-2002	(first entry)
XX		
DT		
XX		
XX	Human full-length TyRS.	
XX		
DE		
XX	Tyrosyl-trna synthetase; TyRS; Rossmann fold nucleotide binding domain;	
KW	vascular endothelial cell function; burn; plastic surgery; abdomen;	
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;	
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;	
KW	dermal ulcer; diabetic ulcer; endothelialization;	
KW	typtophanyl-trna synthetase; tyRS; vascular graft surgery.	
OS	Homo sapiens.	
XX		
PN	WO200174841-A1.	
XX		
PD	11-OCT-2001.	
XX		
PF	21-MAR-2001; 2001WO-US08966.	
XX		
PR	31-MAR-2000; 2000US-193471P.	
XX		
PA	(SCRI) SCRIPPS RES INST.	
XX		
PI	Schimmel P, Wakasugi K.	
XX		
DR	WPI: 2001-626377/72.	
XX	N-PSDB: AAH43602.	

```

QY 158 DINKTFISDLDYMGSSGFEYKNNVYKIOKHYTFNQVKGIFGFTDSDCIGKISFPALQAP 217
   | : | | | | : : : : : : : : : : : : : : | : | | | | | | |
Db 61 DEDKDFITQNSFE-----IKYIEMALPIAKKINFMAKAVEGTEBSKIGMTFFPAIQAP 116
QY 218 SFSNSFPQIFRDRDITQCLIPCAIDDDPYFRMTRVAPRIGYKPKALLHSTFFPALQAG 277
   : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 117 TF-----FERK-----RCILPAAIDDDPYWRLORDEFAESLGYKTAALHSHKFPSTLS 166
QY 278 TKMSASDPNSSIFLTDITAKQITKYNKHAFFSGGRDITIEHRFGNCQDVVSFMTLTFPL 337
   | | | | | | : : | | | : : : | : | : | : | : | : | : | : |
Db 167 GMSASKEPETAALYLDSDPEDVEKKYWKFTLTGGRPPLKEQREKGEPEKCVVFKKLETF 226
QY 338 EDDDLKLEQIRKDY---TSGAMLTGELKRALLEVLOPLAEHQARRREVTDIEVKEFMTPR 394
   | : | | | : : : : : : : : : : : : : : | : | : | : | : | : |
Db 227 EEDDK--KLKEKYACKNCELTCGCKRLLSKIOEFLKEHQRKKR--AEKLVEKFKYTG 283
QY 395 KLS 397
   | :
Db 284 KLA 286

```

RESULT 15

```

Q9HN66 PRELIMINARY; PRT; 380 AA.
AC Q9HN66;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHETASE.
GN TRPS2 OR VNG22326.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahliras G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Sprogon J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isebnarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005109; AAG20355.1;
DR InterPro: IPR001994; Cytidylyltransf.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1; UNKNOWN_1.
KW Aminocacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;

```

Query Match 25.3%; Score 534.5; DB 17; Length 380;
 Best Local Similarity 35.4%; Pred. No. 1.4e-38;
 Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

```

QY 10 AEDDVDPMTVOTSSAKGDYDKLIVRFSSKIDKELINIRERATGQRPHHFLRGIFFS 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 3 ADGNVYTPAIVASDD---LDYEKLLARFAGADELTDQRAFP-----DHLVNRGLFYA 53
QY 70 HRDMNQVLDAVENKRPFLYLTGRGSSSEAMHGHILPIFTKWLQDVENVPLVIQMTDDE 129
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 54 GRDVVDDELTAGQS-----IVTGVGPGS-PMHGHAMVFFARLQDERGARVYVPLSDDE 108
QY 130 KYLWKDLITDQAYGAIVE--NAKDIACGFIDINKTFIF---SPLDYV-GMSSGFEYKNNVYK 184
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

Db 109 KYWEKQTPAET-GDYLRANILRLDLAVGEDELFRIIVVDFRADAVLYPLATAFAGDV--- 164
QY 185 OKHYTFNQVKGIFGFTDSDCIGKISFPALQAPSPNSFPQIFRDRDITQCLIPCAIDOD 244
   : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 165 -RHATLQWVYG-----EPDNWGQAFYPAVQTAHLL--LPQLVHG--BHETLVPLAVDQD 213
QY 245 PYFRMTRVAPRIGYP--KPAIIHSTFFPALQAGQTKMSASDPNSSIFLPTAKOIKTKV 302
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 214 PHVRYSRDYAAKARYPVGKPGALLMQPLPSLAG--PGKMSSS--AGVSTRLTSDPTVREKV 271
QY 303 NKHAFFSGGRDITIEHRQFGNCQDVVSFMYLTFLLEDD--KLEQIRKDYTGAMLTGELK 361
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 272 RTHAATGGRASVEEHRAGGVPAEDVPFQYLSAFEPEDDAELARIEREYRAGDLISGELK 331
QY 362 KALIEVLOPLAEHQARRREVTDIEVKEFMTPRKLSFD 399
   : : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 332 DIAADRTIEFLAAHQRRRALGD--VTBALDAFRLTDD 367

```

Search completed: October 24, 2002, 12:53:05
 Job time : 22.7502 secs

OC Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,
 RA Barrell B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA MEDLINE=9816435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL121862; CAB58393.1;
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 DR Aminoacyl-tRNA synthetase.
 SQ SEQUENCE 490 AA; 54904 MW; 172c68622c511d3d CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;
 Best Local Similarity 38.1%; Pred. No. 2.4e-61;
 Matches 173; Conservative 87; Mismatches 129; Indels 65; Gaps 11;

OY 3 HGPDATEDEDFVDPMTVTQTSANGIDYDKLIVRFGSSKID-----KELINRIERAT- 54
 DB 40 HPRGDAGADVDITPMVVAAKPGQINVDRLTFKAEHMDGARGQMRDVMACRKRMT 99
 OY 55 -----GQR-----HFLRIGIFSRDM 73
 DB 100 TTPPSANAEVAPFVLDAVAGDAALQAAASPAHQORQPPALALHFFHNDIAFSHRL 159
 OY 74 NOVLDAYE-----NKKPFVLYTGKPSSEAMVHGLIPFTFTKLODFVNVPLVIOMTD 128
 DB 160 HKALVLDIEASVKGERSVFLTGKPSGIMHGLVLPMLTKLQDFSLPLVIOITDD 219
 OY 129 EKYLMKDLTDQAVGAV--ENAKDIIACGFDINKTFIFSDLDYMGSSGFKNVVAK 186
 DB 220 EKFLFRDVPPEGAKADELIRNSIKIDIIAFNPFRHFTIFRNTHYMG--DMYPTVLRLR 276
 OY 187 HVPENOVKGIPTDSDICIKISFPALQAPSPNSFPQIFR--DRDIDCLIPCALDOD 244
 DB 277 SMGNAVKHFLGITDSNVKLAEPAPVQAAPCSTAFRVLRLNGDR--PMRCLIPCALDOD 335
 OY 245 PYFRMTDVPAPRIGYPRPALHSTFFPALQGAOTKMSAS--DPNSSIFLTTAKOIKTKVN 303
 DB 336 PFVFLTPAALRLKQLPALHRTKLPALGLEKMSSSAEKGVITLHTDKVKKL- 394
 OY 304 KHAFGSGRDTIEHRQFGNCVDVSPVLTFFLEDDDKLEQIRKDYTSAMLTELKKA 363
 DB 395 RRAFSGGCATLEQMOETGANLELDVAVQYLRFCPPDPTLADVYQRRSGTINSGEVKL 454
 OY 364 LIE-VIOPLIAEHOARKKEYTDELVEKFFMPKRL 396
 DB 455 AADCIIREVLHDMRRERATVTDVVEFCRIRDI 488

RESULT 13
 OY 095295 PRELIMINARY; PRT; 136 AA.
 AC 095295;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE TRYPTOPHANYL-TRNA SYNTHASE (FRAGMENT).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINE;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library";
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z81267; CAB03585.1;
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 FT NON_TER 1
 FT NON_TER 136
 SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

Query Match 32.2%; Score 682; DB 6; Length 136;
 Best Local Similarity 94.1%; Pred. No. 3.6e-52;
 Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 211 PAIQAFSFSFPQIFRDRDIDCLIPCALDODPYFRMTDVPAPRIGYPRPALHSTFF 270
 DB 1 PAIQAFSFSFSFPQIFRDRDIDCLIPCALDODPYFRMTDVPAPRIGYPRPALHSTFF 60
 OY 271 PALQAGTAKMSASDPNSSIFLTDPAKOIKTKVNHAFSGGGRDTEHRQFGNCVDVSVF 330
 DB 61 PALQAGTAKMSAXNPNSIFLTDPAKOIKTKVNHAFSGGGRDTEHRQFGNCVDVSVF 120
 OY 331 MYLTFLEDDDKLEQI 346
 DB 121 MYLTFLEDDDKLEQI 136

RESULT 14
 OY 059584 PRELIMINARY; PRT; 301 AA.
 AC 059584;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE 301AA LONG HYPOTHETICAL TRYPTOPHANYL-TRNA SYNTHETASE.
 GN PH1921.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kusida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000007; BAA31046.1;
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 KW Aminoacyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 301 AA; 35312 MW; 957143E1647CAE1 CRC64;

Query Match 29.6%; Score 626.5; DB 17; Length 301;
 Best Local Similarity 44.2%; Pred. No. 8.3e-47;
 Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

OY 99 MHVGHILPIFTKWLQDFVNVPLVIOMTDDEKYLMD-LTLDQAGDAVENARDIACGF 157
 DB 1 MHIGHILPFATKWLQDFGVNLYIQTDEKFLFKENLTFDDTKRWAYDNIIIDIAVGF 60

Db 201 IWKVVNTQAAIFGFTEDDICKGAAPAVEAACPFASSFPQIFKRRNDICLIPCAIDQ 260
OY 244 DPEFMTDVAAPRIGYKPPALHSTFPALOGAOTKMSASDPNSSIFLTDPKAKOIKTV 302
Db 261 DPEFMTDVAAPRILKASPSLIETFPALGTGATKMSASEPNTICLIPSDAKOIKTV 319

RESULT 10

O976M1 PRELIMINARY: PRT: 406 AA.
ID O976M1
AC O976M1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.
GN ST0169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=11955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamaya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi H.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
KW EMBL, AP000981; BAB5126.1;
KW Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA: 47979 MW: 276CD6506BC94B3 CRC64;

Query Match 43.0%; Score 910.5; DB 17; Length 406;
Best Local Similarity 50.5%; Pred. No. 1,5e-71;
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

OY 12 EDFVDPWTQVOTSSAKG-IDYDKLIVRGSSKIDKELINIRATGQRPHNLRGIGFS 69
Db 28 QDFNTPMEV-----GKVDYDKLIVQFGTKITSELEKIKSIINDELHVMRLRDVFFS 82
OY 70 HDMMQVLDAYENKKRPFLYITGRGSSSEAMHGHLIPFTKWLQDVENVPLVIOMTDE 129
Db 83 HRDLDLVNLKDYDQKGFPLYTGRAPSL-GMHIGHLIPFTKWLQDKEFNVLNLIETDDE 141
OY 130 KILMK-DLTLQDAYGDVAVENAKDIIACGFDINKTFESLDLDYMGSSGFKYKVVKIQKHV 188
Db 142 KMRNRPETLLDQTRQMAYNIIIDITAVGFNPDKTFIFQOTEXI---RNNYPIAIKAKKL 198
OY 189 TFNOYKGIETGFTDSCIGKISFPALQAPSFNSPFOIFRDRTRDIOCLIPCAIDDDPYR 248
Db 199 TSEVAVATGFLTSSNIGITVPAQIAPT-----MEKR---RCLIPAGIDDDPYR 248
OY 249 MRDVAAPRIGYKPPALHSTFPALOGAOTKMSASDPNSSIFLTDPKAKOIKTVNHAS 308
Db 249 LQRDIAESIGYKKAQIHSKFLPPLTGPBGKSSSQPEAIYLLTDPKVERKIMKYAFS 308
OY 309 GGRDTEEHROFGNCDVVSFMYLTFLE-DDDKLEQIRKDYTGAMLTGELKALILEV 367
Db 309 GGQPTTELHRRKTYGNDIDIVSFWMLYMFEPDNKIKKIEEDYRSALLTGELKOLILEK 368
OY 368 LQPLAEHQARRKE 381
Db 369 LNDLEEHRRKEE 382

RESULT 11
O972X0 PRELIMINARY: PRT: 386 AA.

AC O972X0;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).
GN TRPS.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweye M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erasuo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozeza C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL, AE006677; AAK40778.1;
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_1tp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ SEQUENCE 386 AA: 45448 MW: 3F4D8C9B8D219DF CRC64;

Query Match 42.9%; Score 907; DB 17; Length 386;
Best Local Similarity 48.6%; Pred. No. 2.8e-71;
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

OY 8 TEAEEDPDPWTQVOTSSAKG-IDYDKLIVRGSSKIDKELINIRATGQRPHNLRGIGI 66
Db 6 TMDPEFTYPMEV-----GKVDYDKLIVQFGTKITSELEKIKSIINDELHVMRLRDVFFS 59
OY 67 FSHRDMQVLDAYENKKRPFLYITGRGSSSEAMHGHLIPFTKWLQDVENVPLVIOMT 126
Db 60 FSHRDLVLVDLYEKSKGFPLYTGRAPSL-GMHIGHLIPFTKWLQDKEFNANLYIET 118
OY 127 DDEKYLWK-DLTLQDAYGDVAVENAKDIIACGFDINKTFESLDLDYMGSSGFKYKVVKIQ 185
Db 119 DDEKYMNRNPEFTLLDQTRSAVYNDIIDITAVGFNPDKTFIFQOTEXI---RNNYPIYKIA 175
OY 186 KHVTFNOYKGIETGFTDSCIGKISFPALQAPSFNSPFOIFRDRTRDIOCLIPCAIDDDP 245
Db 176 KLTSEVAVATGFLTSSNIGITVPAQIAPT-----MEKR---RCLIPAGIDDDP 225
OY 246 YFRMTDVAAPRIGYKPPALHSTFPALOGAOTKMSASDPNSSIFLTDPKAKOIKTVNHK 305
Db 226 YKRLQDIAESIGYKKAQIHSKFLPPLTGPBGKSSSQPEAIYLLTDPKVERKIMKY 285
OY 306 AFSGRDTEEHROFGNCDVVSFMYLTFLEDDD-KLEQIRKDYTGAMLTGELKALILEV 364
Db 286 AFSGQPTTELHRRKTYGNDIDIVSFWMLYFEEEDNRIKIEEYRSKMLTGELKOLIL 345
OY 365 IEVLQPLAEHQARRKEVTEIVKEMPRKLS 397
Db 346 IDKLNFELEHR-RREAKELVHVKKYDGKLA 377

RESULT 12
O901F5 PRELIMINARY: PRT: 490 AA.
AC O901F5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHETASE.
GN L1063.04.
OS Leishmania major.

QY	64	RGIFSHHDNQAVLDVAENKKRPFLYITGRGSSSEAMNHGHLIPFLFTKMLQDVNVP	LYI	123			
QY <td>64</td> <td>RGIFSHHDNQAVLDVAENKKRPFLYITGRGSSSEAMNHGHLIPFLFTKMLQDVNVP</td> <td>LYI</td> <td>123</td>	64	RGIFSHHDNQAVLDVAENKKRPFLYITGRGSSSEAMNHGHLIPFLFTKMLQDVNVP	LYI	123			
Db <td>67</td> <td>RSVFEEHHDNFENIIDVAERGRDFLYITGRGSSSEALHGLHLPMTFTYLOEAEKVP</td> <td>LYI</td> <td>126</td>	67	RSVFEEHHDNFENIIDVAERGRDFLYITGRGSSSEALHGLHLPMTFTYLOEAEKVP	LYI	126			
QY <td>124</td> <td>QMTDEKXLMQDLTDQAVGVDAVENAKOIIACGPFINTFTFSDUDYVGMSSGFKNVK</td> <td></td> <td>183</td>	124	QMTDEKXLMQDLTDQAVGVDAVENAKOIIACGPFINTFTFSDUDYVGMSSGFKNVK		183			
Db <td>127</td> <td>QUTDEKIMKNLSVEESQRLARENAKDIIACGPDYKTFTFSDPDYVG--GAFYKKNVK</td> <td></td> <td>184</td>	127	QUTDEKIMKNLSVEESQRLARENAKDIIACGPDYKTFTFSDPDYVG--GAFYKKNVK		184			
QY <td>184</td> <td>IOKHENQVNGVIGFETDSDICIGKISFPAIQAAFSFNSPQIFRDRITDCLIPCAIDQ</td> <td></td> <td>243</td>	184	IOKHENQVNGVIGFETDSDICIGKISFPAIQAAFSFNSPQIFRDRITDCLIPCAIDQ		243			
Db <td>185</td> <td>VGKCCTLKAKNGIGFGESEDPILAKISFPVQAVSPFSFPHLEPGKXNCLCLIPCAIDQ</td> <td></td> <td>244</td>	185	VGKCCTLKAKNGIGFGESEDPILAKISFPVQAVSPFSFPHLEPGKXNCLCLIPCAIDQ		244			
QY <td>244</td> <td>DPEYMTFADVAPRIGYKPKPALHSTFPALOGAQTMSASDPNSSIFLTPTAKOIKTVN</td> <td></td> <td>303</td>	244	DPEYMTFADVAPRIGYKPKPALHSTFPALOGAQTMSASDPNSSIFLTPTAKOIKTVN		303			
Db <td>245</td> <td>DPEYMTFADVAPRIGYKPKPALHSTFPALOGENKMSASDPNSAIYVTDKAKIDKNIN</td> <td></td> <td>304</td>	245	DPEYMTFADVAPRIGYKPKPALHSTFPALOGENKMSASDPNSAIYVTDKAKIDKNIN		304			
QY <td>304</td> <td>KHAFSGGSDTEIEHHPQSGNCDDVPSFMYLTFPLEDDKLEQIKRDYVSGMLGELKA</td> <td></td> <td>363</td>	304	KHAFSGGSDTEIEHHPQSGNCDDVPSFMYLTFPLEDDKLEQIKRDYVSGMLGELKA		363			
Db <td>305</td> <td>RYAFSGGSDSTIEKHHELANLEVDPVAKYLSFPLEDDSELEHIFKEYGEGHMLTGEYKRR</td> <td></td> <td>364</td>	305	RYAFSGGSDSTIEKHHELANLEVDPVAKYLSFPLEDDSELEHIFKEYGEGHMLTGEYKRR		364			
QY <td>364</td> <td>LIEVLQPLIAEHQARKEVTEIYKFEFMTPKRLSEDFQ</td> <td></td> <td>401</td>	364	LIEVLQPLIAEHQARKEVTEIYKFEFMTPKRLSEDFQ		401			
Db <td>365</td> <td>LTEVLTEYERHRRARRAAVTDENMDAFNAVAPRLPSMFE</td> <td></td> <td>402</td>	365	LTEVLTEYERHRRARRAAVTDENMDAFNAVAPRLPSMFE		402			
RESULT 9							
ID	0901R2	PRELIMITARY:	PRT:	324 AA.			
AC	0901R2:						
DT	01-MAY-2000 (TREMBLrel. 13, Created)						
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)						
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)						
DE	Y80D3A.1 PROTEIN.						
CN	Y80D3A.1.						
OS	Caenorhabditis elegans.						
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;						
OC	Rhabdilitidae; Pelodierinae; Caenorhabditia.						
OX	NCBI_TaxID=6239;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RA	Harris B.R.;						
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=99069613; PubMed=9851916;						
RA	none;						
RT	"Genome sequence of the nematode C.elegans: A platform for						
RT	investigating biology."						
RT	Science 282:2012-2018(1998).						
DR	EMBL: AL132853; CAB60439.1; "						
DR	InterPro: IPR002305; tRNA-synt_lb.						
DR	InterPro: IPR002306; tRNA-synt_ttp.						
DR	Pfam: PF00579; tRNA-synt_lb; I.						
DR	PRINTS: PRO1039; TRNASYNTHTRP.						
SO	SEQUENCE 324 AA; 36289 MW; 6E687E6E420EC12 CRC64;						
Query Match 48.5%; Score 1027; DB 5; Length 324;							
Best Local Similarity 64.9%; Pred. No. 6.5e-82;							
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;							
QY	4	GGDAAEEDDFDDPTVOTSSAKGIDYDKLIVFGSSKIDKELINIRATGQRPHEFLR		63			
Db	23	GGGVEDEDRVTPEVTTTKATGIDYDKLIVFGCKRLDELTARFRTVGNKSPMLR		82			

Db 83 RGMFAHNDLVALIDRRKGGKPFYLYTERGASSSSGLHGHLPFLFTKWLGEVDFPLVI 142
QY 124 QMTDEKELMKDILTDQAYDAVENAKIILACGPIINTFFISDLDYGMSSGFKNVYK 183
Db 143 QMTDEKELMKDQKVEAKRKARRENMKDIISGEPFTKFTFNPNFY--MCPFEYENLVK 200

Db	208	CVFENQKGI	GFEDSDIL	IGIPRAQA	PAISITFPI	FGNR -	KHCLIPCA	LDPPY	266	
Qy	247	FRMRDVA	PRIGYK	PALLHST	FPFPA	LOQAOT	KMSASD	PNSISFT	DTJAQIKTVKNA	306
Db	267	FRMRDVA	PRIGYK	PALLHST	FPFPA	LOQAOT	KMSASD	PNSISFT	DTJAQIKTVKNA	306
Qy	307	FSGGDIT	IEHRQ	GCNDV	VSFMYL	TFLEDD	KLEQLR	KDYTS	GAMLTGEIKKALTE	366
Db	327	FSGGWVY	VEEHRK	IGGVE	VDVSA	QLKFL	FLEDDAK	LEEVAVAS	KGEMLTGEIKKALTE	386
Qy	367	VLOPIAE	HOARKE	VDDEIK	FEFMP	PRKISF				398
Db	387	TLTPIVE	HOHAKR	LITDE	VDLTK	FEELR	PLKE			418

RESULT 6

ID	0904Y1	PRELIMINARY;	PRT;	430 AA.
AC	0904Y1;			
DT	01-MAY-2000	(TIREMBLrel, 13, Created)		
DT	01-MAY-2000	(TIREMBLrel, 13, Last sequence update)		
DT	01-JUN-2001	(TIREMBLrel, 13, Last annotation update)		
DE	TRYPTOPHANYL-TRNA SYNTHETASE.			
GN	AATS-TRP OR CG9735.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:			
OC	Pterygota: Neoptera: Endopterygota, Diptera; Brachycera: Muscomorpha:			
OC	Ephyraoidea; Drosophilidae; Drosophila.			
XX	NCBI_taxid=7227;			

RN SEQUENCE FROM N.A.
 RP [1]
 RX MEDLINE=99250164; PubMed=10233165;
 RA Seshiah P., Andrew D. J.;
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
 RT the developing Drosophila salivary gland.";
 RL Mol. Biol. Cell 10:1595-1608(1999).
 DR EMBL; AF125156; AAF20166.1; -
 DR Flybase: FBgn0010803; Aats-tdp.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_1tp.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PR01039; tRNAsyntHTRP.
 DR PROSITE; PS00178; AA tRNA-LIGASE_I; 1.
 KW Aminocacyl-tRNA synthetase.
 SQ SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;

Query Match	65.0%;	Score 1374.5;	DB 5;	Length 430;
Best Local Similarity	64.8%;	Pred. No. 3e-112;		
Matches 254;	Conservative 61;	Mismatches 76;	Indels 1;	Gaps 1;

QY	7	A TEAE EEP FVDWMTVOTSSAKIGIDDKILIVRGSSKIKELINLRERATGGRPHHFLTRGI	66
Dd	38	ATAPEEDVDDWMNNASSNDAGVDI DKLIRKGSSKIDEBELAREKITTKGRAHHFIRGM	97
QY	67	FESHDMNQVLDAVENKKPPLYLTGRGPSSSEAMVGHLPFIPTKWLDQVFNNPLVIOMT	126
Dd	98	FESHRLDITLTLREOGKPFLYTGRGPSGLSHGVHLVEIMTKWLQEHREDVEDYLQLT	157
QY	127	DDEKLYMKDLTLDOAYGDAVENADIIACGFDDINKTFISFDLDGMGSSSGFYKNVVKIOK	188
Dd	158	DDEKTYLMKDLEVEDA IKLGRENADIVA IGDGVKTFEINFNLBEFGVCAPMYONMIRIOK	217
QY	187	HATEYNVGKIGIFPDSDPICIKTSIPALOAAPSFNSPOLFRDBETDIQCILICAIODDPY	246
Dd	218	CVTENOVXKGIFGFGSDIIGIKTGPPAQAOPAIISSTPEPIFGNF-KVHCILPCAIODDPY	276
QY	247	FRMRDVAAPRIGYEPALIHSTFPFPALOGAQTOKMSASDPNSSIFRILTJAKOIKTKYKNKA	306
Dd	277	FRMRDVAAPRIGYEPKCALIHSTFPFPALOGAKTKMSASDAQNSAVYLTJDPJOIKNKIKKYA	336
QY	307	FSGGRDITIEHRQGGNGCNDVVYSMWLYTFPLEDDDKLEQLRKDYTSAMLTGELKALKALE	366

Db 337 FSGGVVVEEHHKLCGPEVDVSQQLKEFLEDDAKLEEYNAVYSKSEMLTGELIKLAVE 396

QY 367 VLQPLAEHQARKREVDYEIVKEEMTPRKLSF 398
+ +:: :||| | :|||:: :: + + +
Db 397 TLTPTVEQHQAARKLLTDEVLDKTFELRPPLKF 428

RESULT 7

ID	Q9VHG2	PRELIMINARY;	PRT;	430 AA.
AC	Q9VHG2			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	AATS-TRP PROTEIN.			
GN	AATS-TRP OR CG9735.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephyridiida; Drosophilidae; Drosophila.			
OX	NCBI_taxID=7227;			

RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132;

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.
RA Durbin K.T., Edwards J.G., Edwards C., Foxworth G., Fletcher R.

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Clodagh A., Conn E., Connell T.R., Cuzzocrea S., Curran D., Harris V.

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
BA Hostin D., Houston K.A., Howland T.T., Wei M.-H., Thoenes C.

RA Jalali M., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount C.V., Muehle P., Muehle T., Muehle D.W., Nelson D.T.

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Delaney V., Rittman C., Rittman T., Rittman M.C.

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Crimmins D., Esterhuysen D., Twardochleb D., Webster D., West
J., Williams J., Wilson J., Wood J., Wright J., Young J., Ziegler J.

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RL Science 287:2185-2195(2000).

DR FlyBase; FBgn0010803; Aats-trp.

DR	InterPro; IPR001412; tRNA-synt_I.
DR	InterPro; IPR003306; tRNA-synt_+trn

DR Pfam; PF00579; tRNA-synt_1b; 1.
DB DBINMC; DB01030; tRNACYNTHEBD

DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
SEQUENCE 430 AA; 2028EEFEECC60E070E CPD64.

DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; tRNA-syntHTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244CA42 CRC64;

Query Match 77.0%; Score 1629; DB 11; Length 329;
Best Local Similarity 92.1%; Pred. No. 9.6e-135;
Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 73 MNQVLDAVENKKPFYLTGGRSPSEAMHVGHLIPFTKWLQDVFNPLVLTQMTDDEKYL 132
DB 1 MNQVLDAVENKKPFYLTGGRSPSEAMHVGHLIPFTKWLQDVFNPLVLTQMTDDEKYL 60
QY 133 WKDLTLQAGDAVENAKDITACGPDINKPTFISDLDDYMGSSGFYKNVYIKQHVFNQ 192
DB 61 WKDLTLQAGDAVENAKDITACGPDINKPTFISDLDDYMGSSGFYKNVYIKQHVFNQ 120
QY 193 VKGTFEGTSDSCIGKISFPALQAPSFNSFPQIFRDRDTIQCLIPCAIDDPFRMTRD 252
DB 121 VKGTFEGTSDSCIGKISFPALQAPSFNSFPKLFDRDTIQCLIPCAIDDPFRMTRD 180
QY 253 VAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSSIFLTPAKOIKTKVKNHAFSGRD 312
DB 181 VAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSSIFLTPAKOIKTKVKNHAFSGRD 240
QY 313 TIEHHROFGNCVDVFSFWLTFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 372
DB 241 TIEHHROFGNCVDVFSFWLTFFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 300
QY 373 AEHQARREYTDDELVEKFEFMTPRKLSFDEQ 401
DB 301 AEHQARREYTDDELVEKFEFMTPRKLSFHFQ 329

RESULT 4

070184 PRELIMINARY; PRT; 305 AA.

AC 070184;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TRYPTOPHAN-tRNA SYNTHETASE (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.
NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY; TISSUE=SPLEEN;
RA Yang D., Goto R., Watanabe N., Kobayashi Y.;
RT "Identification and Cloning of Genes Whose Expressions are Elevated
during DNCB-induced Guinea pig Skin Delayed-type Hypersensitivity
Reaction.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
EMBL: AB012222; BAA25288.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; tRNA-syntHTRP.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 1
SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 72.6%; Score 1537; DB 11; Length 305;
Best Local Similarity 94.4%; Pred. No. 1e-126;
Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 97 EAMHVGHLIPFTKWLQDVFNPLVLTQMTDDEKYLWKDLTLQAGDAVENAKDITACG 156
DB 1 EAMHVGHLIPFTKWLQDVFNPLVLTQMTDDEKYLWKDLTLQAGDAVENAKDITACG 60

QY 157 FDIKNTFISDLDDYMGSSGFYKNVYIKQHVFNNOYKIGRFTDSDCIGKISFPALQAA 216
DB 61 FDIKNTFISDLDDYMGSSGFYKNVYIKQHVFNNOYKIGRFTDSDCIGKISFPALQAA 120
QY 217 PSFNSFPQIFRDRDTIQCLIPCAIDDPYFRMTRDVAIRIGYKPKALLHSTFFPALQAA 276
DB 121 PSFNSFPQIFRDRDTIQCLIPCAIDDPYFRMTRDVAIRIGYKPKALLHSTFFPALQAA 180
QY 277 QTKMSADPNSSIFLTPAKOIKTKVKNHAFSGGRDTIEEHROFGNCVDVFSFWLTFF 336
DB 181 QTKMSADPNSSIFLTPAKOIKTKVKNHAFSGGRDTIEEHROFGNCVDVFSFWLTFF 240
QY 337 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIEHQARREYTDDELVEKFEFMTPRKL 396
DB 241 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIEHQARREYTDDELVEKFEFMTPRKL 300
QY 397 SFDFQ 401
DB 301 SFHFQ 305

RESULT 5

0904Y0 PRELIMINARY; PRT; 420 AA.

ID 0904Y0;
AC 0904Y0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRYPTOPHAN-tRNA SYNTHETASE (FRAGMENT).
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99250164; PubMed=10233165;
RA Seshiah P., Andrew D.J.;
RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in
the developing Drosophila salivary gland.";
RL MOL. Biol. Cell 10:1595-1608(1999).
DR EMBL: AF125157; AF20167.1; -
DR FlyBase: FBgn0010803; Aats-trp.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; tRNA-syntHTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 1
SQ SEQUENCE 420 AA; 46975 MW; 4AF7088AFA26AEB6 CRC64;

Query Match 65.0%; Score 1374.5; DB 5; Length 420;
Best Local Similarity 64.8%; Pred. No. 2.9e-112;
Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;

QY 7 ATAEDEDEVPWTVOTSSAKGIDYDKLIVFGSSKIDKELINRERATGGRPHHFLRGI 66
DB 28 ATAEDEDEVPWTVOTSSAKGIDYDKLIVFGSSKIDKELINRERATGGRPHHFLRGI 87
QY 67 FFSHRDMNQVLDAVENKKPFYLTGGRSPSEAMHVGHLIPFTKWLQDVFNPLVLTQMT 126
DB 88 FFSHRDLHTLTLLREQGRKPFYLTGGRSPSGSLHVGHLIPFTKWLQDVFNPLVLTQMT 147
QY 127 DDEKYLWKDLTLQAGDAVENAKDITACGPDINKPTFISDLDDYMGSSGFYKNVYIKQ 166
DB 148 DDEKYLWKDLTLQAGDAVENAKDITACGPDINKPTFISDLDDYMGSSGFYKNVYIKQ 207
QY 187 HATFNQKGIJFGFTSDSCIGKISFPALQAPSFNSFPQIFRDRDTIQCLIPCAIDDPY 246
DB 187 HATFNQKGIJFGFTSDSCIGKISFPALQAPSFNSFPQIFRDRDTIQCLIPCAIDDPY 246

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 21.7502 Seconds
(without alignments)
3189.442 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471
Perfect score: 2116
Sequence: 1 SNHGPDATAEEDFVDPMTV.....VTDEIVKEMTPRKLSFDQ 401

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPTRMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oranella:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1957	92.5	475	11	Q9DC65
2	1957	92.5	481	11	Q9J558
3	1629	77.0	329	11	Q9DBR9
4	1537	72.6	305	11	Q70184
5	1374.5	65.0	420	5	Q9U4Y0
6	1374.5	65.0	430	5	Q9U4Y1
7	1370.5	64.8	430	5	Q9VHG2
8	1301	61.5	402	10	Q9SR15
9	1027	48.5	324	5	Q9UIR2
10	910.5	42.9	406	17	Q976M1
11	907	42.9	386	17	Q972X0
12	795.5	37.6	490	5	Q9UIF5
13	682	32.2	136	6	Q9S295
14	626.5	29.6	301	17	Q95984
15	534.5	25.3	380	17	Q9NH66
16	310	14.7	111	5	Q95YL8

17	299.5	14.2	426	17	Q978Y8	Q978Y8 thermoplasma
18	296.5	14.0	136	6	Q9TS88	Q9TS88 bos taurus
19	294	13.9	157	5	Q9US33	Q9US33 trypanosoma
20	286	13.5	109	4	Q9UD15	Q9UD15 homo sapien
21	269.5	12.7	513	17	Q9HN83	Q9HN83 halobacteri
22	268.5	12.7	426	17	Q9HIW5	Q9HIW5 thermoplasma
23	188	8.9	364	17	Q9YA64	Q9YA64 aeropyrum p
24	185.5	8.8	341	16	Q97N42	Q97N42 streptococc
25	182.5	8.6	340	16	Q99XH4	Q99XH4 streptococc
26	181.5	8.6	331	17	Q979Z1	Q979Z1 thermoplasma
27	174.5	8.2	341	16	Q9CDD1	Q9CDD1 lactococcus
28	167	7.9	351	16	Q9RVD6	Q9RVD6 delnococcus
29	164.5	7.8	895	10	Q9SGN2	Q9SGN2 arabidopsis
30	163	7.7	460	10	P93018	P93018 arabidopsis
31	158.5	7.5	375	17	Q9V027	Q9V027 pyrococcus
32	154.5	7.3	102	1	007119	007119 halobacteri
33	152.5	7.2	682	5	Q9NDB8	Q9NDB8 leishmania
34	149.5	7.1	327	17	Q9HN62	Q9HN62 halobacteri
35	149.5	7.1	408	10	P93363	P93363 nicotiana t
36	146	6.9	375	17	Q58739	Q58739 pyrococcus
37	145	6.9	332	17	Q9HKR3	Q9HKR3 thermoplasma
38	142.5	6.7	528	11	Q91WQ3	Q91WQ3 mus musculu
39	139	6.6	294	17	Q96TV3	Q96TV3 sulfolobus
40	135	6.4	365	16	Q9KNV7	Q9KNV7 vibrio chol
41	134.5	6.4	525	5	Q9VW60	Q9VW60 drosophila
42	130.5	6.2	419	16	Q92B81	Q92B81 listeria in
43	129	6.1	339	2	Q9KZ47	Q9KZ47 streptococ
44	124	5.9	344	16	Q9AC05	Q9AC05 caulobacter
45	123	5.8	347	10	082313	082313 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q9DC65	PRELIMINARY;	PRT;	475 AA.
AC	Q9DC65:			
DT	01-JUN-2001 (TREMURel. 17, Created)			
DT	01-JUN-2001 (TREMURel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMURel. 19, Last annotation update)			
DE	ADULT MALE LUNG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,			
DE	CLONE:1200002C07, FULL INSERT SEQUENCE.			
GN	WARS.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=LUNG;			
RX	MEDLINE=21085660; Pubmed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaudo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK004541; BAB23357.1; -.			


```

QY 219 FSNFPOIFRDRDIOCLIPCAIDODPYFRMTRDVAPPRTGYKPPALLSTFFPAL-OGAQ 277
      : | | : | | : | | : | | : | | : | |
Db 176 -----LDEQFLDVDCQFG-GVDQRKIFVLAENLPISLGKKRAHLMNPMVPGLAQCG- 226
      : | | : | | : | | : | | : | | : | |
QY 278 TKMSASDPNSSIFLTDPAKQIKTVNKKHAFESGGRDTIEH-----ROFGG 322
      : | | : | | : | | : | | : | | : | |
Db 227 -KMSASDPNSKIDLEEKQYKKKINSAFCSPG--NVEENGLISFVOYVIAPIQELKFGT 283
      : | | : | | : | | : | | : | | : | |
QY 323 NCDVDVSEFMYLTFELEDDBK-----LEQIRKDYTSGAMLTGELK---KALIEYLO 369
      : | | : | | : | | : | | : | | : | |
Db 284 N-----HFEFIDRPEKFGPITYKSFEEEMKLAPEEKLSPDLKIGVADAINELE 335
      : | | : | | : | | : | | : | | : | |
QY 370 PLIAHOARRKRYTDEIVKEF--MTPRK 395
      : | | : | | : | | : | | : | | : | |
Db 336 P-IRQEFANNKEFQEAASEKGYVATPOK 362

```

Search completed: October 24, 2002, 12:51:44
 Job time : 8.35218 secs

KW Ligase; ATP-binding.
 FT SITE 91 100 "HIGH" REGION.
 FT SITE 275 279 "KMSKS" REGION.
 SQ SEQUENCE 395 AA; 44910 MW; E656AE8B76C5FDE9 CRC64;
 Query Match 57.28; Score 1210; DB 1; Length 395;
 Best Local Similarity 59.68; Pred. No. 1.2e-92;
 Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 11 EEDVDWPTVOTST-----SAGIDVYDKLIVRGSSKIDKELINLRERATGGRPHFLRGI 66
 DB 4 EEOIVTWMVKGSTVDEEKGIDERLIVGTRKITPEOLEPEREKLTKKPHLLRGA 63
 QY 67 FFSHRDNQVLDAYENKRPFLYTGSGPSSSEAMHGHILPIFTKWLQDVFNVLQMT 126
 DB 64 FFSHRDMDLIDREQKKPFLYTGSGPSSSMHGHIMPEKFWLQDVFNVLQMT 123
 QY 127 DDEKYLKMD-LTLDQAYGDAVENAKDIIAGCGDINKTPIESDLDMGSSGFYKVVQIQ 185
 DB 124 DDEKFLFKOGVLEDCORFAENAKDIIAVGFDPKTFIEFNSTYVG--GAFYQNVVRIA 181
 QY 186 KHVFENVKGIFFGFTSDSCIGKISFPALQAAFSNSPQIFRRTDIOCLIPALIDDP 245
 DB 182 KCTIANKSKACFGFTSDSICGKIFASIOAAPSSSPHIFENAKDIPCLIPALIDDP 241
 QY 246 YFRMTDVARIRIGYKPKALLSTFFPALQGAQTMMASDPNSIFLDTAKOIKTKYKHK 305
 DB 242 YFRITRVSGLKPKKALLHSRFPALQGSKMSASKDSALFPMIDTPKIKINKINRH 301
 QY 306 AFSGGRDITIEHRQFGNCVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALI 365
 DB 302 AFSGSGATITIEHRKGNPNVDVAYQYLSFLLDDEKLLQLYNTRYKAGLTSGEMKECI 361
 QY 366 EYIQLPLAEHQARKREYTDVYKFEPM--PKLSF 398
 DB 362 KLLQGFVSDFOARSKVDENTLDMFMGSKRLEW 395

RESULT 6
 SYNC_YEAST STANDARD; PRT; 432 AA.
 ID STWC_YEAST
 AC 012109;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--
 tRNA ligase) (TPRS).
 GN WRS1 OR YOL097C OR HRE432.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96076631; Pubmed=7502582;
 RA Vandenbol M., Durand P., Portetalle D., Hilger F.;
 RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
 including the Tyl-H3 retrotransposon, the sufi(+) frameshift
 RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Tyr-Ia and a
 RT delta element.";
 RL Yeast 11:1069-1075(1995).
 RN [2]
 RP FUNCTION
 RX MEDLINE=97197969; Pubmed=75046085;
 RA John T.R., Ghosh M., Johnson J.D.;
 RT "Identification and expression of the Saccharomyces cerevisiae
 RT cytoplasmic tryptophanyl-tRNA synthetase gene.";
 RL Yeast 13:37-41(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: Z48149; CAAB8164.1; -;
 DR EMBL: Z74839; CA99110.1; -;
 DR SGD: S0005457; WRS1.
 DR InterPro: IPR002305; tRNA-syn_1b.
 DR InterPro: IPR001412; tRNA-syn_1.
 DR InterPro: IPR002306; tRNA-syn_trp.
 DR Pfam: PF00579; tRNA-syn_1b; 1.
 DR PRINTS: PR01039; TRNASYNTHTRP.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_1; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT SITE 111 120 "HIGH" REGION.
 FT SITE 295 299 "KMSKS" REGION.
 SQ SEQUENCE 432 AA; 49350 MW; C408F169737E9736 CRC64;

Query Match 55.08; Score 1163; DB 1; Length 432;
 Best Local Similarity 54.88; Pred. No. 1e-88;
 Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEEDFVDPWY-----QTSSAKGIDVYDKLIVRGSSKIDKELINLRERATGGRPHHF 61
 DB 19 STDVKEQVYVPMDEGGVDGGRQANDIDYKLLKQFETKPVNETTLKRFQVYGRPHHF 78
 QY 62 LRRCIFSSHDMQVLDAYENKRPFLYTGSGPSSSEAMHGHILPIFTKWLQDVFNVL 121
 DB 79 LRKLFPSERDFKILDLLEGKRPFLYTGSGPSSSMHGHIMPEFTKWLQDVFNVL 138
 QY 122 VIQMTDEKYLKMD-LTLDQAYGDAVENAKDIIAGCGDINKTPIESDLDMGSSGFYK 180
 DB 139 VIELTDEKFLFKHKLITINDVKNFARENAKDIIAVGFDPKTFIEFSDLYMG--GAFYET 196
 QY 181 VVKIQKIVTNQVKGITGFSDSCIGKISFPALQAAFSNSPQIFRRTDIOCLIPCA 240
 DB 197 VWRVSQITGSTAKAVGFDSDSCIGKIFHASIOATAFSPSPFNVLGDPKTPCLIPCA 256
 QY 241 IDDDPYFRMTDVARIRIGYKPKALLSTFFPALQGAQTMMASDPNSIFLDTAKOIKT 300
 DB 257 IDDDPYFRVCRDVAADKTKSKPALLHSRFPALQGSITKMSASDDTIAITMTDPKOIK 316
 QY 301 KVNKAFSGGRDITIEHRQFGNCVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360
 DB 317 KINKYARSGGVSADLHRELGNPNVDVAYQYLSFDDVFLKECYDKYKSGELLSGEM 376
 QY 361 KKAILEVLOPLAEHQARKREYTDVYKFEPMTRKL 396
 DB 377 KLCIETLQELQVAKAFQERRAQVDETLKPMVPRKL 412

RESULT 7
 ID STW_PYRAB STANDARD; PRT; 385 AA.
 AC 09UY11;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TPRS).
 GN TRPS OR PAB1111.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=ORSAY;
 RC Helbig R.;

```

|||||
Db 315 IDDPYFMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKT 374
QY 301 KVNKHAFGSGRDTIEEHROFGNCVDVSMYLTFFLEDDDKLEDIRKDYSGAMLTGEL 360
Db 375 KVNKHAFGSGRDTIEEHROFGNCVDVSMYLTFFLEDDDKLEDIRKDYSGAMLTGEL 434
QY 361 KKALEVLQPLAEHOARKREVTDELVEKFMTPRKLSDFQ 401
Db 435 KKELDIVQLPLAEHOARKREVTDEMVKFMTPRQLCFHYQ 475

```

RESULT 4

```

SYM_MOUSE STANDARD: PRT: 481 AA.
ID SYM_MOUSE

```

```

AC P32921:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN MARS OR MRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95018226; PubMed-7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
EMBRYONIC STEM CELLS.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69656; CAA49347.1; -
DR EMBL: X69657; CAA49348.1; -
DR PIR: S31461; S31461.
DR PIR: S31462; S31462.
DR MGD: MGI:104630; MARS.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PF00458; WHEP-TRS; 1.
DR PRINTS: PF01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.

```

```

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VAAPSLIC 476 481 MISSING (IN MAJOR ISOFORM).
FT SEQUENCE 481 AA: 54282 MW: 80544520C08074F52 CAC64;

```

```

Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 9,7e-153;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;

```

```

QY 2 NHGDATEAEEDVDPWTVOTSSAKGIDYDKLIYFGSSKIDKELINIERATGQRPNNF 61
Db 76 NCDSDATKASEDFVDPWTVRTSSAKGIDYDKLIYQPGSSKIDKELINIERATGQRPNNF 135
QY 62 LRRGFFSHRDMNOVLDAENKKRPYLYTGRGSSSEAMHVLIPFTKMLQOVNPL 121
Db 136 LRRGFFSHRDMNOVLDAENKKRPYLYTGRGSSSEAMHVLIPFTKMLQOVNPL 195
QY 122 VIQMTDEKYLKMDLTLDQAGDAVENAKDIIACGFDINTKFTIPSDIDYMGSSGFYKNV 181
Db 196 VIQMSDDEKYLKMDLTLEQASVYENAKDIIACGFDINTKFTIPSDIDYMGSSGFYKNV 255
QY 182 VKIQKHTFENVKGIFFGTSDDCIGKISFPAIQAPSFNSFPQIFRDRDIDQCLIPCAI 241
Db 256 VKIQKHTFENVKGIFFGTSDDCIGKISFPAIQAPSFNSFPQIFRDRDIDQCLIPCAI 315
QY 242 DDDPYFMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKTK 301
Db 316 DDDPYFMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKSK 375
QY 302 VNKHAFGSGRDTIEEHROFGNCVDVSMYLTFFLEDDDKLEDIRKDYSGAMLTGELK 361
Db 376 VNKHAFGSGRDTIEEHROFGNCVDVSMYLTFFLEDDDKLEDIRKDYSGAMLTGELK 435
QY 362 KKALEVLQPLAEHOARKREVTDELVEKFMTPRKLSDFQ 401
Db 436 KKELDIVQLPLAEHOARKREVTDEMVKFMTPRQLCFHYQ 475

```

RESULT 5

```

SYM_SCHPO STANDARD: PRT: 395 AA.
ID SYM_SCHPO

```

```

AC Q09692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
ligase) (TrpRS).
GN SPAC2F7.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxId=4896;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: Z50142; CAA90500.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X53918; CAA37872.1; -
 DR EMBL: X52113; CAA36356.1; -
 DR EMBL: M74074; AAA30799.1; -
 DR PIR: A40279; YRBO.
 DR PIR: S14540; S14540.
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR PRINTS: PRO1039; TRNASYNTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 24 69 WHEP-TRS.
 FT SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
 FT SITE 169 178 "HIGH" REGION.
 FT SITE 353 357 "KMSKS" REGION.
 FT CONFLICT 17 17 L -> M (IN REF. 2).
 SQ SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 95.98; Score 2028.5; DB 1; Length 475;
 Best Local Similarity 95.88; Pred. No. 3.1e-160;
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SNHPPDAEEDVDVDPWTQVTSAGKIDYDKLVREGSSKIDKELNRIERATGQRPNH 60
 DB 76 SGEGLDTEADEDFVDVDPWTQVTSAGKIDYDKLVREGSSKIDKELNRIERATGQRPNH 135
 QY 61 FLRRGIFFSHRDMNQVDAENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVFNVP 120
 DB 136 FLRRGIFFSHRDMNQVDAENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVFNVP 195
 QY 121 LVIQMTDEKYLKDLTLDDAYGDAVENAKDIACGFDINKTFIFSLDYMGSSGEYKN 180
 DB 196 LVIQMTDEKYLKDLTLDDAYGDAVENAKD-ITCGFDINKTFIFSLDYMGSSGEYKN 254
 QY 181 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIRDRDVOCLIPCA 240
 DB 255 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIRDRDVOCLIPCA 314
 QY 241 IDDPYFRMTRDVAPRIGYPRKPLALHSTFPFALOGAQTKKMSASPNSIFLDTAKOIKT 300
 DB 315 IDDPYFRMTRDVAPRIGYPRKPLALHSTFPFALOGAQTKKMSASPNSIFLDTAKOIKT 374
 QY 301 KVNKHAFFSGGKDTLEHRGEGNCDVDSFMYLTFPLEDDKLEQIRKDYSGAMLTGEL 360
 DB 375 KVNKHAFFSGGKDTLEHRGEGNCDVDSFMYLTFPLEDDKLEQIRKDYSGAMLTGEL 434
 QY 361 KKALIEVLQPLIAEHOARKREVTDEIVKEEPTPKLSFDPQ 401
 DB 435 KKALIEVLQPLIAEHOARKREVTDEIVKEEPTPKLSYDFQ 475

RESULT 3
 SYM_RABIT
 ID SYM_RABIT STANDARD; PRT; 475 AA.
 AC P23612; Q28607;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TrpRS).
 GN MARS.
 OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90239043; PubMed=2185472;
 RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
 RT "Cloning and expression of a mammalian peptide chain release factor
 RT with sequence similarity to tryptophanyl-tRNA synthetases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
 RN [2]
 RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
 RX MEDLINE=94009008; PubMed=8404867;
 RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
 RA McLaughan K.K., Kisseliev L.L., Tate W.P., Haenni A.-L.;
 RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
 RT synthetase are distinct proteins.";
 RL EMBO J. 12:4013-4019(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
 CC RELEASE FACTOR (ERF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33460; AAA31246.1; ALT-SEO.
 DR EMBL: U02595; AAB60257.1; -
 DR PIR: A35904; YWRBR.
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam: PF00579; tRNA-synt_1b; 1.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR PRINTS: PRO1039; TRNASYNTRP.
 DR PROSITE: PS00178; AA-TRNA-LIGASE_1; 1.
 DR PROSITE: PS00762; WHEP-TRS; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 23 68 WHEP-TRS.
 FT SITE 168 177 "HIGH" REGION.
 FT SITE 353 357 "KMSKS" REGION.
 SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718F45DC4 CRC64;

Query Match 92.78; Score 1961.5; DB 1; Length 475;
 Best Local Similarity 91.58; Pred. No. 1.1e-154;
 Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 2 NHG-PDAEAEEDVDVDPWTQVTSAGKIDYDKLVREGSSKIDKELNRIERATGQRPNH 60
 DB 75 SHGPEAVADKEDFVDVDPWTQVTSAGKIDYDKLVREGSSKIDKELNRIERATGQRPNH 134
 QY 61 FLRRGIFFSHRDMNQVDAENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVFNVP 120
 DB 135 FLRRGIFFSHRDMNQVDAENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVFNVP 194
 QY 121 LVIQMTDEKYLKDLTLDDAYGDAVENAKDIACGFDINKTFIFSLDYMGSSGEYKN 180
 DB 195 LVQVMSDEKYLKDLTLDDAYGYGTLENARDIACGFDVKNKTFIFSLDYMGSSGEYKN 254
 QY 181 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIRDRDVOCLIPCA 240
 DB 255 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIFHGADIOCLIPCA 314
 QY 241 IDDPYFRMTRDVAPRIGYPRKPLALHSTFPFALOGAQTKKMSASPNSIFLDTAKOIKT 300

RA Frolowa L.Y., Grigorieva A.Y., Sudomocina M.A., Kisselev L.L.;
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-
 RT response elements and exon-intron organization.";
 RL Gene 128:237-245(1993).
 RN [7]
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.
 RC TISSUE-Keratinocytes; PubMed=1286667;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Cells J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [8]
 RP FUNCTION.
 RX MEDLINE=92225128; PubMed=1373391;
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;
 RT "An interferon-induced protein with release factor activity is a
 RT tryptophanyl-tRNA synthetase.";
 RL FEBS Lett. 300:162-166(1992).
 CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1 SUBUNIT: HOMODIMER.
 CC -1 INDUCTION: BY INTERFERON GAMMA.
 CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1 SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC
 DR EMBL; M77804; AAA67324.1; -;
 DR EMBL; X59892; CAA42545.1; -;
 DR EMBL; M61715; AAA61298.1; -;
 DR EMBL; M62570; CAA44450.1; -;
 DR EMBL; S82905; AAB39381.1; -;
 DR EMBL; X67920; CAB94198.1; -;
 DR EMBL; X67921; CAB94198.1; JOINED.
 DR EMBL; X67922; CAB94198.1; JOINED.
 DR EMBL; X67923; CAB94199.1; -;
 DR EMBL; X67924; CAB94199.1; JOINED.
 DR EMBL; X67925; CAB94199.1; JOINED.
 DR EMBL; X67926; CAB94199.1; JOINED.
 DR EMBL; X67927; CAB94199.1; JOINED.
 DR EMBL; X67928; CAB94199.1; JOINED.
 DR PIR; A41706; A41706.
 DR PIR; A41633; A41633.
 DR PIR; JH0533; JH0533.
 DR AATHUS/Genent-2DPAGE; 3524; IEF.
 DR P1C1-2DPAGE; P23381; -;
 DR MIM; 191050; -;
 DR InterPro: IPR000738; WHEP-TRS.
 DR InterPro: IPR002305; tRNA-synt_1b.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002306; tRNA-synt_trp.
 DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR Pfam; PF00458; WHEP-TRS; 1.
 DR PRINTS; PRO1039; TRNASYNTHTRP.
 DR PROSITE; PS00178; AA-TRNA_LIGASE_L; 1.
 DR PROSITE; PS00762; WHEP-TRS; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 19 64 WHEP-TRS.
 FT SITE 164 173 "HIGH" REGION.
 FT SITE 349 353 "KMSKS" REGION.
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).
 FT CONFLICT 424 424 A -> R (IN REF. 4).
 SQ SEQUENCE 471 AA; 53165 MW; E9634449053A0D0 CRC64;

Query Match 99.3%; Score 2101; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 3,1e-166;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SNHGPDATAEEDFPDPTVOTSSAKGIDYDKLIVRFSSKIDKELINRIEATGQRPH 60
 DB 71 SNHGPDATAEEDFPDPTVOTSSAKGIDYDKLIVRFSSKIDKELINRIEATGQRPH 130
 OY 61 FLRGIFFSHRDMNOVLAVENKRPFLYTGPGSSAMHVGHLIPFTFTKLODFVNP 120
 DB 131 FLRGIFFSHRDMNOVLAVENKRPFLYTGPGSSAMHVGHLIPFTFTKLODFVNP 190
 OY 121 LVIQMTDEKYLWKDLTLDQAYDAVENAKDIIACGPDINKTFISDDYMGSSGFYKN 180
 DB 191 LVIQMTDEKYLWKDLTLDQAYDAVENAKDIIACGPDINKTFISDDYMGSSGFYKN 250
 OY 181 VVKIKKHTFNVKRIIFGFTSDCIGKISFPAIQAAPSFNSFPDIFDRDIOCLICA 240
 DB 251 VVKIKKHTFNVKRIIFGFTSDCIGKISFPAIQAAPSFNSFPDIFDRDIOCLICA 310
 OY 241 IDDDYFRTNRDVARPRIGYPRKALHSTFFPALQAGOTKMSASDPNSSIFLDTAKOIKT 300
 DB 311 IDDDYFRTNRDVARPRIGYPRKALHSTFFPALQAGOTKMSASDPNSSIFLDTAKOIKT 370
 OY 301 KVNKAFFSGGRDTEIENHROFGNCDDVVSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 360
 DB 371 KVNKAFFSGGRDTEIENHROFGNCDDVVSFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 430
 OY 361 KKALIEVLOPLAEHQARKEVTDEIVKEFMTPRKLSDFQ 401
 DB 431 KKALIEVLOPLAEHQARKEVTDEIVKEFMTPRKLSDFQ 471
 RESULT 2
 SYM_BOVIN
 ID SYM_BOVIN STANDARD; PRT; 475 AA.
 AC P17248;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
 DE (TPRS).
 GN WARS.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE-Relina;
 RX MEDLINE=91329348; PubMed=1907847;
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,
 RA Gandar J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to
 RT prokaryotic synthetases but near identity with mammalian peptide
 RT chain release factor";
 RL Biochemistry 30:7809-7817(1991).
 RN [2]
 RP SEQUENCE OF 17-475 FROM N.A.
 RC TISSUE-Pancreas;
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., La Bounesse B.,
 RA Labouesse J., Bonnet J.;
 RT Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
 CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
 CC diphosphate + L-tryptophanyl-tRNA(Trp).
 CC -1 SUBUNIT: HOMODIMER.
 CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1 SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 7.35218 Seconds
(without alignments)
2111.829 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471
Perfect score: 2116
Sequence: 1 SNHGPDATGEEDFVDPWTV.....VTDEIVKEFWTPRKLSDFQ 401

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	2101	99.3	471	1	SYW_HUMAN
2	2028.5	95.9	475	1	SYW_BOVIN
3	1961.5	92.7	475	1	SYW_RABIT
4	1938	91.6	481	1	SYW_MOUSE
5	1210	57.2	395	1	SYW_SCHPO
6	1163	55.0	432	1	SYWC_YEAST
7	803	37.9	385	1	SYW_PYRAB
8	409.5	19.4	370	1	SYW_METJA
9	397.5	18.8	364	1	SYW_METTH
10	386	18.2	134	1	SYW_ENCCU
11	370.5	17.5	420	1	SYW_ARCFU
12	353	16.7	374	1	SYW_MERPE
13	192	9.1	323	1	SYW_ARCFU
14	175.5	8.3	341	1	SYW_CLOLO
15	172	8.1	394	1	SYWC_YEAST
16	170.5	8.1	366	1	SYW_SULSO
17	169	8.0	395	1	SYW_AOUAE
18	166.5	7.9	346	1	SYW_CHLUR
19	163.5	7.7	353	1	SYW_BORBU
20	159.5	7.5	337	1	SYW_WREPA
21	158	7.5	346	1	SYW_CHLMU
22	152	7.2	344	1	SYW_CHLPP
23	146	6.9	326	1	SYW_HELPP
24	138	6.5	319	1	SYW_METTH
25	136.5	6.5	528	1	SYW_BOVIN
26	135.5	6.4	328	1	SYW_THEMA
27	133	6.3	328	1	SYW_BACST
28	132.5	6.3	343	1	SYW_MYCLE
29	132.5	6.3	528	1	SYW_HUMAN
30	130	6.1	326	1	SYW_HELPP
31	129	6.1	401	1	SYWC_SCHPO
32	128.5	6.1	379	1	SYWM_YEAST
33	128	6.0	334	1	SYW_HAEIN

34	124.5	5.9	347	1	SYW_MYCLE	P47372 mycoplasma
35	122	5.8	350	1	SYWM_CAEBL	P46579 caenorhabdi
36	118.5	5.6	343	1	SYW_CLOAB	G971d6 clostridium
37	116.5	5.5	330	1	SYW_BACSU	P21656 bacillus su
38	112.5	5.3	335	1	SYW_BUCAT	P57602 buchnera ap
39	110	5.3	306	1	SYW_METJA	Q57834 methanococc
40	109.5	5.2	336	1	SYW_MCTVU	Q53386 mycobacteri
41	107	5.1	330	1	SYW_RICPR	G92d76 rickettsia
42	106.5	5.0	334	1	SYW_ECOLI	P00954 escherichia
43	105.5	5.0	333	1	SYW_PASMO	P57956 pasteurella
44	105	5.0	426	1	SYW_MYCLE	Q49900 mycobacteri
45	104	4.9	337	1	SYW_SYNY3	P73655 synechocyst

ALIGNMENTS

RESULT 1	SYW_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381: P78535; Q9UDL3:			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)			
DE	(TrpRS) (IFP53) (hmRS).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutaria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92105071; PubMed=1761529;			
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92107982; PubMed=1763065;			
RA	Flecker J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RL	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92112058; PubMed=1765274;			
RA	Prolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovjeva O.L.,			
RA	Kisselev L.L.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RL	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92164636; PubMed=1537332;			
RA	Buitt U., Flohr T., Boeltger E.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RL	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RX	MEDLINE=96319944; PubMed=8724762;			
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-t			
RT	RNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RL	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE-Sperm;			
RX	MEDLINE=93292992; PubMed=7685728;			

[illegible]

RESULT 11
F64476
tryptophan--rRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii
N:Alternate names: tryptophanyl-rRNA synthetase
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

R. Bulst, C.J.; White, O.; Olsen, G.U.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reisch, C.I.; Overbeek, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodok, A.; Hannon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

A. Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A. Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.
A. Reference number: A64300; MUID: 96337999

A:Accession: F64476
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-370 (BMT)

A:Accession: U00096, GB:U00096.1, PID:g1592065, T:1
 A:Cross-References: GB:U67582, GB:L77117, NID:g1592064, PID:AAB99425.1, PID:g1592065, T:1
 C:Genetics:
 A:Map position: FOR1375885-1376997
 A:Start codon: GTG
 C:Superfamily: mammalian tryptophan--tRNA ligase: amino acid--tRNA ligase repeat homologous
 Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match	19.4%	Score 409.5;	DB 2;	Length 370;
Best Local Similarity	30.9%	Pred. No. 1.1e-25;		
Matches 119; Conservative	72;	Mismatches 153;	Indels 41;	Gaps 16;

0y 17 PWTVOSSAKGIDYDKLIVRFEGSSKIDKELINRIERATGQRPHPFLRGIFESHDMNOV 76
+ + : : : : : : : : : : : : : : : :
db 8 PW-ETPAV-IDYKRTMEFGFKPIVDVLGDJKEE-----HHFFRNIIILGHDFERI 57
+ + : : : : : : : : : : : : : : : :

```

QY 77 LDAYENKKPPLYLTGRCPSSEAMHVGHILPIFTKWLQ---DVFNPLVIQMTDDEKYL 13
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 YDAKNNKEPAAVSGMPSGK-MHEGHNKVVLDLLEKYOKYTDNINPI---ADLEAWA 111

```

Qy 134 KDLTLDQAYGBAV-ENAKDIICGFIDINKTFIFESBLDYMGSSGPKNVKTI-QKHVTFN 19
::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 113 RNMEFEETKEALINEYTNVTATIGIDPEFTNNVYSKYCKV-----KDIALITISKPTMMS 167

```

Oy 192 QVKGJFGFTSDCIGKISFPAIQAPSPSNSFPQIFRDRT--DIOCLIPCAIDQDPYERM 249
      :|:|:|      ||:|:|:|      ||:|:|      :|:|:|:|:|:|:|
Db 168 EMKVEICEKEKENTGICVADITQVADDT--HPQIDENKCNEDKPLUUDVCTTQDNNHT 225

```

QY 250 TRDYAPR---IGYPKRALHSTFFPALOGAQTKMSASDPNSSIFLTDTAKQTKTKVNKHA 306
				: : :					:		:
				: : :					:		:
				: : :					:		:

```

307 FSGCHDTIEHRQFG--NCDDVSEFWLTFPEEDDKIEQIRKDYTSGAMLTGELKKA 36
      :|::||::||::|| | | ::| : || : || : || : || : ||

```

DU 202 NAGREITLERNKIGVPEECVIELELH-HLLDDNELELHIOCRSGELIGCKRM 338
QY 364 LIEVLQPLIAEHQARRKEVTDEIVK 388
| : : : : : ||

Db 340 AYERVVEFLKDLKEKREQAKEIAVK :

RESULT 12
E69131
trycphan--tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum, (strain D
N;Alternate names: trycphanyl-tRNA synthetase
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence,revision 05-Dec-1997 #text,change 22-Oct-1999

Rasmussen, D.R.: Doucette-Stamm, L.A.: Deloughery, C.: Lee, H.: Dubois, J.: Alredge, T.: Qiu, D.: Spadafora, R.: Viscarte, R.: Wang, Y.: Wierowski, J.: Gibson, R.: Jivanli, K.: Church, G.M.: Daniels, C.J.: Mao, J.: Rice, P.: Noelling, J.: Reeve, J.N.: *Bacteriol.* 179, 7135-7155, 1997

A: Accession: E69131
A: Reference number: A69000; MUID:98037514
A: Molecule type: DNA
A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA
A: Residues: 1-364 <MTH>
A: Cross-references: GB:AE000812; GH:AE000666; NID:g2621296; PIDN:AAE84757.1; PID:g2622
A: Experimental source: strain Delta H
A: Genetics:

A:Gene: MTH251
A:Start codon: TTG
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo
C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match	18.8%;	Score 397.5;	DB 2;	Length 364;
Best Local Similarity	27.9%;	Pred. No. 1e-24;		
Matches 107;	Conservative 76;	Mismatches 153;	Indels 47	

```
QY      15 VDPWVTQTSSAKAGIDYKDLIVRGSSKIDKELINRIERATGQRPHFLRGIFFSHRDMN 74
       :||| ||| :| | | | | :| :| :| :| | | | | | | | | | | | | | | |
Db      2 IDPW----GSAK-LEYYQDLIENGVRRP--SEVLDEV-----PEPSMLMRGGIIFGHRIYE 50
```

QY 75 QVLDAYENKKRPFYLTYGRGPSSSEAMHGILIPFTIKWLQDVFNVPVLVIQMDDKEYLWK 138
:: : : : : : : : : : : : : : : : : : :
Db 51 RIISAKKGGEEFVAVTGMPSPGF-MHIGKMIVDQLRW-YDRMGAEIIFIPIDMEAVSAR 106

Qy 135 DLTLDQAYGAVEN-AKDIACGFDINK-----TFIFSLDLYMGSSGFFKNNVK 183

Db 109 GVDGDSRRIATIEFYTAGYTAGIDIEKNIHVVYLOSENIMVEDIAYV----- 156

QY 184 IQKHVFNQVKIGFGFTDSDCIKISIPAIQAAPSFNSFPQIFRDRTDICLIPCAlDQ 243
:
: | | : : | | | | : : | | :
157 TACKNENFPAIVCECSEMAHVAITVOVEDI UNOI DEL CCDD - - DIVVIDENDO 313

244 DPFRRTRDVAPRI----GYPKALLHSFFFPALOGAQTMSASDPNSSIFLDTAKOIK 299

300 TVNKHAFSGRDTIEHRQFGNCDDVVSFMYLTFFLE--DDDKLEQIRKDYTSGLMTG 358

Qy 359 ELKALIEVLPLIAEHQARKE 381
| | : : | : : :

DO 331 ECONNIALEF LNFFEEELSVAKEN 333

RESULT 13

143806

```
C:Accession: M3806
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Species: Encephalitozoon cuniculi
C:trypTophan--tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)
```

A1Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein

A:Reference number: Z22093; MUID:98217683
A:Accession: T43806
A:Status: preliminary; translated from GB/EMBL
A:Molecule type: DNA

Db 346 IDKLNFLLEHR-RRREAKELVHFXYDGKLA 377

RESULT 8

C75020 tryptophanyl-tRNA synthetase (trps) PAB1111 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: C75020

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome, stru

A:Reference number: A75001

A:Accession: C75020

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <RAW>

A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:GAB50601.1; PID:g545911

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: trps, PAB1111

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match

Best Local Similarity 45.3%; Score 803; DB 2; Length 385;

Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

Db 12 EDF-VDMVOTSSAKIDVDKILVREGSSKIDKELINRERATGPRPHFLRGIFPSH 70

Db 3 EDFVTFMEVEGV---VDYKNKLIHFHTSPLELLEKTVLETKSELPLFFRRKFFFSH 58

Db 71 RDMQVLDVAVENKRPFLYTGSGSEAMHGHILPFETKWLQDVNPVLIQMTDDEK 130

Db 59 RDVQKVLQDVEEGKFLYTGSGSG-PMHGHILPFPATKWLQDEKGVNLIYITDDEK 117

Db 131 YLMKD-LTLDQAYGDAVENAKDIACGDIKTFIFSDLDGMGSGFYKNVAKIQKHVT 189

Db 118 FLFRENLTPEETKHWAVENILDIIVGDPDKTFIFQNSE---TKYEMAIPIAKIN 173

Db 190 FNOVKGIFGTFDSCICIGISFPAIQAPSSNSFPQIFRDTDIQCLIPCAIDDDPYFRM 249

Db 174 FSKAKAVFGFTEQSKIGMIFPAIQAPTF-----FEKR---RCLIPALIDDDPYWRL 223

Db 250 TRDVAIPGVPKPAHLSTPEPALOGATKMSASDPNSITLTPAKOIKTKVKNHAFSG 309

Db 224 QRDPAESLGYIKTAHSKFSPLTSLSGKMSASKPETALVLTSPDEVEKKVKKFALTG 283

Db 310 GRDIEHROFGNCDVDVSEFMYLTFLEDDDKLEQIRKDY---TSGAMLTGELKALIE 366

Db 284 GRPLTKQREKGEPEKCVVFKMLEIFFEEDDK--KLKERYACKNGELTGCCKRYLIS 341

Db 367 VLQPLIAHQARKEVDIETK 388

Db 342 KIOFLKEHQRRKKAKEQIEK 363

RESULT 9

G71206 tryptophan--tRNA ligase (EC 6.1.1.2) - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: G71206

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Onikubo, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Koshida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; MUID:98344137

A:Accession: G71206

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-301 <RAW>

A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BA31046.1; PID:g3258363

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by Genba

C:Genetics:

A:Gene: PH1921

C:Superfamily: yeast tyrosine--tRNA ligase

C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match

Best Local Similarity 44.2%; Score 626.5; DB 2; Length 301;

Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

Db 99 MHVGLIPFTTKWLQDVNPVLIQMTDDEKYLMD-LTLDQAYGDAVENAKDIACGF 157

Db 1 MHIGHIIPFATKWLQDEKGVNLIQITDDEKFLFKEKNTLPEDTRMAVNDIILVAVG 60

Db 158 DINKTFISDLDYMGNSGFGYKNVYKQKHTFNOVGIFGFTSDICIGISFPAIQAP 217

Db 61 DPDKTFIFQNSE---TKYEMAIPIAKKINFSMAAVFETQSKIGMIFPAIQAP 116

Db 218 SFSNSFPQIFRDTDIQCLIPCAIDDDPYFRMTDVAIPRIQPKPALHSTFPALOGAQ 277

Db 117 TF-----FEKR---RCLIPALIDDDPYWRLQDPAESLGYIKTAHSKFSVSLTSL 166

Db 278 TRMSASDPNSITLTPAKOIKTKVKNHAFSGRDTIEHRQFGNCDVDVSEMYLTFEL 337

Db 167 GKMSASKPETALVLTSPDEVEKKVKKFMTLGRPTLKEQREKGEPEKCVVFKMLEIFE 226

Db 338 EDDDKLEQIRKDY---TSGAMLTGELKALIEVLQPLIAHQARKEVDIETK 394

Db 227 EEDDK--KLKERYACKNGELTGCCKRYLISKIOEFLKEHQRRRK-AEKLVEKFTYG 283

Db 395 KLS 397

Db 284 KLA 286

RESULT 10

G84373 tryptophanyl-tRNA synthetase (imported) - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: G84373

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky

; Leitshuser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, G.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: G84373

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <STO>

A:Cross-references: GB:AE004437; NID:g10581646; PIDN:AAG20355.1; GSPDB:GN00138

C:Genetics:

A:Gene: trpS2

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match

Best Local Similarity 25.3%; Score 534.5; DB 2; Length 380;

Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

Db 10 AEEFVDPMTVOTSSAGIDYDKLIVFGSSKIDKELINRERATGPRPHFLRGIFPS 69

Db 3 ADGNDVVPVAESD---LDEKTLARFGDELTDGDRARP-----DHELVNRGLTYA 53

Db 70 HRDQNVLDVAVENKRPFLYTGSGSEAMHGHILPFIFKWLQDVNPVLIQMTDDE 129

Db 54 GRVDVDFLTAGEQS---IYTGVPSPG-PHILGHAMFYFARRLQDDEFGARVYPLSDDE 108

Db 130 KYLMKDLTLDQAYGDAVE-NAKDIACGFDINKTFIF---SDLDY-GMSSGFGYKNVYK 184

Db 109 KYWFKDQTPAET--GDYLRANRLDLLAVGDPPELIRIVVDRDADVLVPLATAFAGDV--- 164

Db 124 DDEKFLFGQVSLDQCQFARENAKDIIAVGDFPKKTFIFMNSTYVG--GAFYQVWVRLA 181
QY 186 KHVFNQVKGIFGFTDSDCTGKISPAIOAAPSPNSFPQIFRDRDTIOCLIPCAIDDDP 245
Db 182 KCIATNOSKACGFFDSDSICKIFHFAISIOAASFSSSPHFNGAKDIPCLIPCAIDDDP 241
QY 246 YFRMRDVAAPRIGYKPKPALHSTFPFALOGAOTKMSADSPNSIFLTPDIAKQIKTKVNHK 305
Db 242 YFRLLRDVSGRLKFKKPPALLHSRFPALOGPQSKMSASKDSAIITMTDIPKIKKINH 301
QY 306 AFSGGRTDIEHRQFGNCDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKAL 365
Db 302 AFSGGATIEIHRKGGNPDVAVOYLSPFLDDEKLEQIYNTYKAGTLSTGEMKGBCI 361
QY 366 EYLDPLIAEHQARKEVYDEIVKEFMT--PKLSF 398
Db 362 KLLQGFVSDFOAARSKVDATLDMFMDSRKLEW 395

RESULT 6

S51901

tryptophan--tRNA ligase (EC 6.1.1.2) [similarity] - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein HRE332; protein 00792; tryptophanyl-tRNA synthetase
C:Species: Saccharomyces cerevisiae
C:Date: 05-May-1995 #sequence_revision 03-Aug-1995 #text_change 20-Jun-2000
C:Accession: S51901; S59177; S66793
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
Submitted to the EMBL Data Library, January 1995
A:Description: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including and a Delta.
A:Reference number: S51848

A:Accession: S51901

A:Molecule type: DNA

A:Residues: 1-432 <VAM>

A:Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256

R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.

Yeast 11, 1069-1075, 1995

A:Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the a delta element.

A:Reference number: S59156; MUID:96076631

A:Accession: S59177

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-432 <VAM>

A:Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

R:Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.

Submitted to the Protein Sequence Database, July 1996

A:Reference number: S66791

A:Accession: S66793

A:Molecule type: DNA

A:Residues: 1-432 <DNR>

A:Cross-references: EMBL:Z74839; NID:g1419947; PIDN:CAA9110.1; PID:g1419948; GSPDB:GN0C

A:Experimental source: strain S28C

C:Genetics:

A:Gene: WRS1; MIPS:YOL097C

A:Map position: 15L

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

C:Keywords: ligase

F:117-120/Region: ATP-binding motif (HXGH)

Query Match 55.0%; Score 1163; DB 2; Length 432;
Best Local Similarity 54.8%; Pred. No. 3.8e-87;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEDFVDPWTV-----QTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPNHF 61
Db 19 STDVEQVQVTPDWDEGVGDEQGRQAINIDYDKLIKQFGTPVNEETLKRKQVGTGREPHNF 78
QY 62 LARGGFSSRDNNQVLDAYENKKPFIYVGRGSSSEAMVGLIFIFIFKMLQDVFNVP 121
Db 79 LKKGIFSERDFTKILLDYEGKPFPLTYGKGPSSDSMHLGHIPIVFTKWLQEVFVDP 138

QY 122 VIOMTDEKYLWK-DLTLDOAYGDAVENAKDIIACGFDINKTFIFPSDLDYMGSSGFYKN 180
Db 139 VIELTDEKFLPKKHLITINDKFNARENAKDIIAVGDFPKKTFIFPSDLYQYMG--GAFYET 196
QY 181 VKIQRVHTVNOVKGIFGFTDSDCTGKISPAIOAAPSPNSFPQIFRDRDTIOCLIPCA 240
Db 197 VYRVSRLQTGSAKAVAFGFDNSDCTGKIFHFAISIOAASFSSSPHFVNLGIPDKTIPCLIPCA 256
QY 241 IDDDYFRMTDVAAPRIGYKPKPALHSTFPFALOGAOTKMSADSPNSIFLTPDIAKQIKT 300
Db 257 IDDDYFRMTDVAAPRIGYKPKPALHSTFPFALOGAOTKMSADSPNSIFLTPDIAKQIKT 300
QY 301 KVNKAFFSGGRDTEIHRQFGNCDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360
Db 317 KINKAFFSGGVYKADLHRELGNPNPDVAVOYLSFRKDDVFLKCYKYSNGELLSEGM 376
QY 361 KKALEVLQPLIAEHQARKEVYDEIVKEFMTPRKL 396
Db 377 KKLCTETLQEFVKAQERRAQVDEETLDFKFWPHKL 412

RESULT 7

C90190

tryptophanyl-tRNA synthetase (trps) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: C90190
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Anayez, M.J.; Ch
Jong, T.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
Submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139

A:Accession: C90190

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-386 <KUR>

A:Cross-references: GB:AB06641; NID:g13813608; PIDN:AAK40778.1; GSPDB:GN0155

C:Genetics:

A:Gene: trps

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 42.9%; Score 907; DB 2; Length 386;
Best Local Similarity 48.6%; Pred. No. 2.7e-66;
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

QY 8 TEAEDFVDPWTVQTSAGK-IDYDKLIVRGSSKIDKELINRIERATGQRPNHF 66
Db 6 TWPDEFVTPWEV-----GKQVYDKLIVQFQTKITELKQIRKNLAGDL-HVMLRRNV 59
QY 67 FFSHRDNNQVLDAYENKKPFIYVGRGSSSEAMVGLIFIFIFKMLQDVFNVP 126
Db 60 FFSHRDLDVLDVYKRSKGFELYTGRAPSL-GMHIGHLIPFTWLDKFKFANLYIET 118
QY 127 DDEKYLWK-DLTLDOAYGDAVENAKDIIACGFDINKTFIFPSDLDYMGSSGFYKNVYK 185
Db 119 DDEKFMRNDEFLDQTRSAKYNNIILDIAGVGNPKTFIFQDTEVI---RNMYPTTVXIA 175
QY 186 KHVFNQVKGIFGFTDSDCTGKISPAIOAAPSPNSFPQIFRDRDTIOCLIPCAIDDDP 245
Db 176 KKLTFSEVRAFPGLDASSNIGLIFYPALQIAPT-----MEKR--RCLIPACIDDDP 225
QY 241 IDDDYFRMTDVAAPRIGYKPKPALHSTFPFALOGAOTKMSADSPNSIFLTPDIAKQIKT 300
Db 226 YRMRDVAAPRIGYKPKPALHSTFPFALOGAOTKMSADSPNSIFLTPDIAKQIKT 305
QY 301 KVNKAFFSGGRDTEIHRQFGNCDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKAL 364
Db 286 AFSGGATIEIHRKGGNPDVAVOYLSPFLDDEKLEQIYNTYKAGTLSTGEMKGBCI 361
QY 361 KKALEVLQPLIAEHQARKEVYDEIVKEFMTPRKL 397
Db 377 KKLCTETLQEFVKAQERRAQVDEETLDFKFWPHKL 412

A:Residues: 1-212, 'GD', 215-471 <FRO2>
A:Cross-references: GB:MB1715; NID:g340367; PIDN:AAA61298.1; PID:g340368
A:Experimental source: fibroblast
C:Genetics:
A:Gene: GDB:WARS; IIP53
A:Cross-references: GDB:119632; OMIM:191050
A:Map position: 14q23-14q31
A:Introns: 33/2; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homology
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
C:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match	99.3%	Score 2101;	DB 1;	Length 471;
Best Local Similarity	99.5%	Pred. No. 1e-163;		
Matches 399;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	1	SNHGDA	TEAEED	FVDPMT	VTQTS	SAKGDYDK	LIVRGSSK	IDKEL	INRIR	AGQR	RHH	60																									
Db	71	SNHGDA	TEAEED	FVDPMT	VTQTS	SAKGDYDK	LIVRGSSK	IDKEL	INRIR	AGQR	RHH	130																									
Qy	61	FLRRGIF	FSHSDM	NOVD	LAYEN	KKPEY	LYTGR	GPSS	EMAH	GHLIP	FTKMLD	VFWP	120																								
Db	131	FLRRGIF	FSHSDM	NOVD	LAYEN	KKPEY	LYTGR	GPSS	EMAH	GHLIP	FTKMLD	VFWP	190																								
Qy	121	LVIQMT	DEKYL	WMDL	TLDO	AYGDA	VENAK	DI	IACG	FD	INTFT	IFSD	L	180																							
Db	191	LVIQMT	DEKYL	WMDL	TLDO	AYGDA	VENAK	DI	IACG	FD	INTFT	IFSD	L	250																							
Qy	181	VVKIQKH	TFNQ	YK	GI	ISF	FTSD	CGIK	IS	FPAL	QAAP	SF	NS	FO	IF	DR	DI	O	CL	IP	CA	240															
Db	251	VVKIQKH	TFNQ	YK	GI	ISF	FTSD	CGIK	IS	FPAL	QAAP	SF	NS	FO	IF	DR	DI	O	CL	IP	CA	310															
Qy	241	IDDPY	FPMTR	DVAP	PRIGY	PKP	PAL	HS	TFF	PAL	OGA	Q	KMS	AS	DP	NS	I	F	L	T	D	RA	KO	IK	300												
Db	311	IDDPY	FPMTR	DVAP	PRIGY	PKP	PAL	HS	TFF	PAL	OGA	Q	KMS	AS	DP	NS	I	F	L	T	D	RA	KO	IK	370												
Qy	301	KVNKHA	FSGGR	D	TI	IEH	RQ	FG	NC	D	V	S	F	M	L	T	F	F	E	D	D	K	L	RQ	IR	KD	Y	T	S	G	A	M	T	G	E	L	360
Db	371	KVNKHA	FSGGR	D	TI	IEH	RQ	FG	NC	D	V	S	F	M	L	T	F	F	E	D	D	K	L	RQ	IR	KD	Y	T	S	G	A	M	T	G	E	L	430
Qy	361	KKALIE	V	I	O	P	L	A	E	H	O	A	R	K	E	V	T	D	I	A	K	E	F	M	T	P	K	L	S	D	F	Q	401				
Db	431	KKALIE	V	I	O	P	L	A	E	H	O	A	R	K	E	V	T	D	I	A	K	E	F	M	T	P	K	L	S	D	F	Q	471				

RESULT 2
YWBO
tryptophan--trna ligase (EC 6.1.1.2) [validated] - bovine

A:Name: tryptophanyl-tRNA synthetase
 C:Species: Bos primigenius laurus (cattle)
 C:Date: 30-Sep-1992 #sequence,revision 30-Sep-1992 #text_change 21-Jul-2000
 C:Accession: A40279; JN0354; S10460; S14540
 R:Garret, M.; Pajot, B.; Trezequet, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti
 Biochemistry 30, 7809-7817, 1991
 A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic S
 A:Reference number: A40279; MUID:91329348
 A:Accession: A40279
 A:Molecule type: mRNA
 A:Residues: 1-475 <GAR>
 A:Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799
 A:Experimental source: pancreas
 A:Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue
 A:Note: part of this sequence was confirmed by protein sequencing
 R:Zarganova, T.A.; Kovaleva, G.K.; Favorova, O.O.; Aeylina, H.B.; Telezhinskaya, I.H.
 Bioorg. Khim. 15, 1307-1311, 1989
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca
 A:Reference number: JN0354; MUID:90211408
 A:Accession: JN0354
 A:Molecule type: protein
 A:Residues: 112-124;282-287, 'N', 288, 'E', 289-292, 'Q', 293-294, 'IR', 336-353;423-441,443-449
 A:Experimental source: liver
 A:Note: this paper is in Russian
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homolog

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
F;24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match	95.9%;	Score 2028.5;	DB 1;	Length 475;
Best Local Similarity	95.8%;	Pred. No. 8.7e-158;		
Matches 384;	Conservative 8;	Mismatches 8;	Indels 1;	Gaps 1;

QY	1	SNHGDAEAEDEPDDMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQPHH	60
Db	76	SGEGDIAEAEDDEPDDMTVOTSSAKGIDYDKLIVRGSSKIDELVNRITERATGQPHR	1355
QY	61	FLRRCIFFSHRDMNOVLDAVENKRPFLYLTGRGSEAMHGHLIPRIFPKWLODFNVN	120
Db	136	FLRRCIFFSHRDMNOVLDAVENKRPFLYLTGRGSEAMHGHLIPRIFPKWLODFNVN	135
QY	121	LVIOQTFDEKYLMDLTLDOAYGDAVENAKDIACGFDINKTFIESDLDMGSSGEYKN	180
Db	196	LVIOQTFDEKYLMDLTLDOAYGDAVENAKD-ITCGFDINKTFIESDLDMGSSGEYKN	254
QY	181	VVKI0KHTFN0VNGIFSEFTSD0ICKISPAI0AASFSNSPFOIRDRDTDLOCLIPCA	240
Db	255	VVKI0KHTFN0VNGIFSEFTSD0ICKISPAI0AASFSNSPFOIRDRDTDLOCLIPCA	314
QY	241	ID0PYFEMTNDVAPRIGPKPALHSTFFPALG0AOTKMSAS0PNSISFLTPAKOIKT	300
Db	315	ID0PYFEMTNDVAPRIGPKPALHSTFFPALG0AOTKMSAS0PNSISFLTPAKOIKT	374
QY	301	KVNKAHFSGGRDJIEHNQFSGNCDVDVNSFWYLFLEDDDKLEQIRKDDYTSGLMTGEL	360
Db	375	KVNKAHFSGGRDJIEHNQFSGNCDVDVNSFWYLFLEDDDKLEQIRKDDYTSGLMTGEL	434
QY	361	KKALIEVLOPLIAEHQARRKEVTDIELYKEFMTPKKLSDFDQ	401
Db	435	KKALIEVLOPLIAEHQARRKEVTDIELYKEFMTPKKLSDFDQ	475

RESULT 3 YWRBPR

tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)
 C:Date: 30-Sep-1992 #sequence_revision 13-Feb-1998 #text_change 26-May-2000
 C:Accession: A35904, S37396
 R:Lee, C.C.; Craigen, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequen-
 A:Reference number: A35904; MUID:90235043

A:Accession: A35904
A:Molecule type: mRNA
A:Residues: 1-475 <LEB>
A:Cross-references: CB:M33460
R:Ritola, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegon, G.; McCaughan, K.
EMBO J. 12, 4013-4019, 1993
A>Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase
A:Reference number: S37396; MUID:94009008
A:Accession: S37396
A:Molecule type: mRNA
A:Residues: 166-177 <PRO>
C:Genetics:
A:Gene: WRS
C:Complex: homodimer [validated, MUID:94009008]
C:Function:
A:Description: EC 6.1.1.2 [validated, MUID:94009008]: catalyzes the ATP-dependent for-
A>Note: mammalian WRS (tryptophanyl-tRNA synthetases) and eRF (polypeptide chain rele-
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo-
C:Keywords: aminocyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protea-
F:23-68/Domain: amino acid--tRNA ligase repeat homology <ATL>
E:174-177/Region: ATP-binding motif (HXGH)

Query Match	92.0%;	Score 1946.5;	DB 1;	Length 475;
Best Local Similarity	90.8%;	Pred. No. 4.3e-151;		
Matches 364;	Conservative 21;	Mismatches 15;	Indels 1;	Gaps 1;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 13.479 Seconds
(without alignments)
2858.658 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116

Sequence: 1 SNHGPDATGEDEFDVDPMTV.....VTDEIVKEFMTPRKLSDFQ 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR_71:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

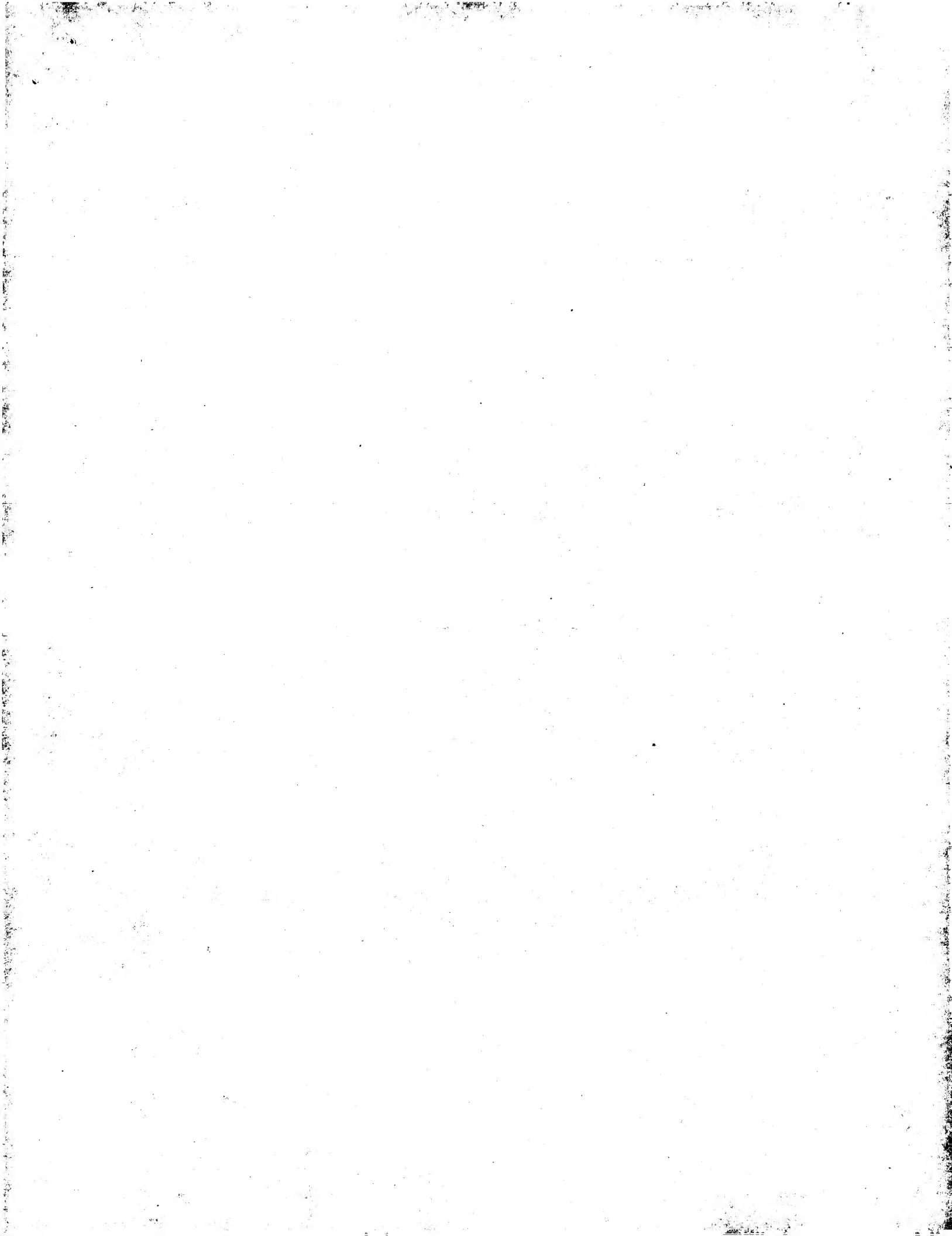
Result No.	Score	Query Match	Length	DB	ID	Description
1	2101	99.3	471	1	A41706	tyrptophan--trna 1
2	2028.5	95.9	475	1	YWBO	tyrptophan--trna 1
3	1946.5	92.0	475	1	YWRBR	tyrptophan--trna 1
4	1938	91.6	481	2	S50053	tyrptophan--trna 1
5	1210	57.2	395	2	S58157	hypothetical prote
6	1163	55.0	432	2	S51901	tyrptophan--trna 1
7	907	42.9	386	2	C90190	tyrptophanyl--trna
8	803	37.9	385	2	C75020	tyrptophanyl--trna
9	626.5	29.6	301	2	G71206	tyrptophan--trna 1
10	534.5	25.3	380	2	G84373	tyrptophanyl--trna
11	409.5	19.4	370	2	F64476	tyrptophan--trna 1
12	397.5	18.8	364	2	E69131	tyrptophan--trna 1
13	386	18.2	134	2	T43806	tyrptophan--trna 1
14	370.5	17.5	420	2	E69461	tyrptophanyl--trna
15	353	16.7	374	2	D72477	probable tyrtopha
16	269.5	12.7	513	2	F84371	tyrptophanyl--trna
17	192	9.1	323	2	H69346	tyrosyl--trna synth
18	188	8.9	364	2	E72512	probable tyrosyl-t
19	185.5	8.8	341	2	D95260	tyrptophanyl--trna
20	185.5	8.8	341	2	G98125	tyrptophan--trna 1
21	174.5	8.2	341	2	B86633	tyrptophan--trna 1
22	172	8.1	394	2	A45999	tyrptophan--trna 1
23	170.5	8.1	366	2	S75410	tyrptophan--trna 1
24	169	8.0	395	2	H70385	tyrptophan--trna 1
25	167	7.9	351	2	E75438	tyrptophanyl--trna
26	166.5	7.9	346	2	B71496	tyrptophan--trna 1
27	164.5	7.8	895	2	A86410	protein F3M18.22
28	163.5	7.7	353	2	E70100	tyrptophan--trna 1
29	163	7.7	460	2	C84750	probable tyrosyl-t

30	159.5	7.5	337	2	F71300	tyrptophan--trna 1
31	158.5	7.5	375	2	B75072	tyrosyl--trna synth
32	158	7.5	346	2	C81654	tyrptophanyl--trna
33	154.5	7.3	102	2	T44994	probable tyrtopha
34	152	7.2	344	2	H86590	tyrptophanyl trna
35	152	7.2	344	2	C72034	tyrptophan--trna 1
36	149.5	7.1	327	2	C84374	tyrosyl--trna synth
37	149.5	7.1	408	2	T03741	probable tyrosine-
38	146	6.9	339	2	E64676	tyrptophan--trna 1
39	146	6.9	375	2	F71093	tyrosine--trna 1
40	145	6.9	337	2	A11066	tyrptophan--trna 1
41	138	6.5	319	2	H69102	tyrosine--trna 1
42	135.5	6.4	328	2	C72370	tyrptophan--trna 1
43	135	6.4	365	2	E82052	tyrptophanyl--trna
44	132.5	6.3	343	2	S73024	tyrptophan--trna 1
45	132	6.2	328	1	YWBSF	tyrptophan--trna 1

ALIGNMENTS

RESULT 1

A41706
tyrptophan--trna ligase (EC 6.1.1.2) [similarity] - human
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor h
C:Species: Homo sapiens (man)
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000
C:Accession: A41633; A41706; S19246; JN0676; JN0533; S26287
R:Flecker, J.; Rasmussen, H.H.; Justesen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9
A:Reference number: A41633; MUID:92107982
A:Accession: A41633
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <FILE>
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 24245-24248, 1991
A:Title: Interferon induces tryptophanyl-trna synthetase expression in human fibrobls
A:Reference number: A41706; MUID:92105071
A:Accession: A41706
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:M7804; NID:9184656; PIDN:AAA67324.1; PID:9184657
R:Buwilt, U.; Flohr, T.; Boettger, E.C.
EMBO J. 11, 489-496, 1992
A:Title: Molecular cloning and characterization of an interferon induced human cDNA w
A:Reference number: S19246; MUID:92164636
A:Accession: S19246
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423, 'R', 425-471 <BUW>
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709
A:Note: 213-Ser and 214-Tyr were also found
R:Frolova, L.Y.; Grigorjeva, A.Y.; Sudomina, M.A.; Kisselev, L.L.
Gene 128, 237-245, 1993
A:Title: The human gene encoding tryptophanyl-trna synthetase: Interferon-response el
A:Reference number: JN0676; MUID:93292992
A:Accession: JN0676
A:Molecule type: DNA
A:Residues: 1-141;182-471 <FRO>
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X6792
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB
A:Note: the authors translated the codon GGC for residue 55 as Cys and GAG for residu
A:Note: this translation is not annotated in GenBank entries HSWRSX1A, HSWRSX1B, HSWR
R:Frolova, L.Y.; Sudomina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kisselev, L.L.
Gene 109, 291-296, 1991
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr
A:Reference number: JN0533; MUID:92112058
A:Accession: JN0533
A:Molecule type: mRNA



Fri Oct 25 09:23:11 2002

us-09-813-718-10_copy_71_471.ra

Page 9

Search completed: October 24, 2002, 12:54:34
Job time : 10.8029 secs

Db 117 QEIADMKERYQRGIGLVKTKRYLLEILERLGPRIERRIEFANDM 162

RESULT 14

US-08-415-593-45
Sequence 45, Application US/08415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Miajun
APPLICANT: Politis-Virk, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumococcal Aminoacyl tRNA
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CP194-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 370 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-415-593-45

Query Match 4.9%; Score 104.5; DB 2; Length 370;

Best Local Similarity 20.2%; Pred. No. 0.0041;
Matches 52; Conservative 44; Mismatches 114; Indels 47; Gaps 10;

QY 71 RDMNOVDAYENK-----PPLYTGRGSPSEAMHGHLPFT-----FTK----- 111
Db 17 RGLDEVLGAERLRKLEERDLKLTWGTSPGKRP-HCGYFVPMIKIADFLKAEVEVTLIFA 75
QY 112 ---WLDQVFNPIVYQMDDEKYLKMDLTLDQAYGDAVENKDIICAGDFINKEFIISDL 168
Db 76 DIHAFDLNKLAPIDLVKRAKRYEFTIAILKSGVSTEKLPVAGSSYQJSSKTCMDN- 134
QY 169 DYMGSSGFYNNVYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAPSPSNSFPQIFR 228
Db 135 -----FRICITVFEHDAKKAGAEVVKQVENSLSGLLPGMOA-----LDEEYL 178
QY 229 DRTDIOCLIPCAIDODPYFRMTRDVAPRIQYKPKALLHSTFPALOGAQTMSAS-DPNS 287
Db 179 D-SDAQF---GGVDQRKLTFAEKRYLPMGLKRLIHLMSPIGLAGG--KMSASGNNEN 232
QY 288 SIFLDTFAKQIKTKYVNK 304
Db 233 KIDILDAEYVKKKINK 249

RESULT 15

US-08-844-054-2
Sequence 2, Application US/08844054
Patent No. 6071892
GENERAL INFORMATION:
APPLICANT: Hodgson, John
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6071892el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19046
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,054
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9608001.5
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31454
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-844-054-2

Query Match 4.5%; Score 95; DB 3; Length 418;

Best Local Similarity 21.2%; Pred. No. 0.059;
Matches 67; Conservative 50; Mismatches 129; Indels 70; Gaps 15;

QY 64 RGIFESHDMNOVDAYENKRPFLYTGSGPSSAMHGHLPFTFKWLDQVFNVPYVI 123
Db 10 RGLIFQTTDEALKRKALEGQVSY-YTGYDPTASLHGHVAILTSRRQLAGHKRYAL 68
QY 124 -----QMTDEKRYLKMDELTDQAYGDAVENKDIICAGFDI---NKTFIISDL 168
Db 69 VGATGILGIDPSFDAERSLQTKDTVD---GWVKSIOGQLSRFLDENGKNKAVMVNNY 124
QY 169 DYMGSSGFYNNVYKIOKHTFNQVKGIFGFTSDCIGKISFPALQAPSPSNSFPQIFR 228
Db 125 DMFG-SISFIDFLRDIDIGYFTVNYM-----MSKESYK---RIETGISYTERAYOIMQ 173
QY 229 D-----RTDIOCLIPCAIDODPYFRMT-----RDVAPRIQYKPKALLHSTFPALOGA 276
Db 174 GYDFEVLNODHNVTLOIG-GSDQMGNTAGTELLRRADKRG-----HYITVPILITDA 225
QY 277 QTKMSASDPNSSIFLPTDAKQIKTKYVNKHAFSSGRDITIEHRQFGN-CDVD-VSPMYLT 334
Db 226 TGRKFKGSEGNVWLNPE-----KTSPEYMYQPMWVMDADAVRLFILIF 269
QY 335 FELEDDDKLEQIRKDY 350
Db 270 TFLSLDE-IEDIRKQF 284

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-100-4

Query Match
Best Local Similarity 5.1%; Score 108; DB 3; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSTFFPALOGAQTMSA 282
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLEVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEQARKREVTDEI 386
DB 117 QEIDMKERYQGGIGVKTKRYLLELRELGPTRERRIEFKAMD 162

RESULT 12
US-09-183-134-2
Sequence 2, Application US/09183134
Patent No. 6165759
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 2
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/183,134
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-183-134-2

Query Match
Best Local Similarity 5.1%; Score 108; DB 4; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSTFFPALOGAQTMSA 282
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLEVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLEVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEQARKREVTDEI 386
DB 117 QEIDMKERYQGGIGVKTKRYLLELRELGPTRERRIEFKAMD 162

RESULT 13
US-09-492-581-4
Sequence 4, Application US/09492581
Patent No. 6346409
GENERAL INFORMATION:
APPLICANT: Gentry, Danlle
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6346409el trps
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,581
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-492-581-4

Query Match
Best Local Similarity 5.1%; Score 108; DB 4; Length 197;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSTFFPALOGAQTMSA 282
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLEVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEQARKREVTDEI 386
DB 117 QEIDMKERYQGGIGVKTKRYLLELRELGPTRERRIEFKAMD 162


```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-743-130A-2

Query Match
Best Local Similarity 20.7%; Score 162.5; DB 2; Length 409;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAVENK-KPEFLYTGSGPSSEAMHGHLPFI-----FTKMLQDV-----F 117
D 27 IKDYLEKRNRPVKIYWGAPTKGP-HCGYFVPMIKLHFLKAGCEVTVLADLHAFLLDM 85
QY 118 NPELYIOMTDEKYLKMLTLDDQAYGDAVENAKDIIACGFDINKTFIFSDIDYMGSSGF 177
D 86 KAPLEVYKRAKYEEFVVKALKLSINPIERLKFVYGSSYOKGDDYV--MDLEKLSNIV 142
QY 178 YKNVVK-----IOKHVFNOYKGTGFTDSDICIKISFPALQAAFSFNSFPQIFRDTD 232
D 143 SQNDARRAGADVKKVQANPLISGLI-----IPLMQA-----IDEEHLG 180
QY 233 IOCLIPCAIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAGQTKMSASDPNSSIFLT 292
D 181 VDAQRG-GVDQRKIFVLAEMNLPISGYKKRAHLMNPMVPGI-GGCGKMSASDPNSKIDII 238
QY 293 DTAKOIKTKVNHAFSGG--RDT-----IIEHRQFCGN 323
D 239 EEPKVVKKKVSAYCAPGELKDNGLIAFLIEVYIOPIALKTVGEGAFKLDIDRPERKYG- 297
QY 324 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALL-----EVLQPLIAHQARR 379
D 298 ---PLST-----DSIDQLKADFYVDKLPPLDKLVADKINELLAPRAEFESS- 343
QY 380 KEVTDEIVKEFMTPRK 395
D 344 -----EEFOVAQK 351

RESULT 6
US-08-705-868-4
; Sequence 4, Application US/08705868
; Patent No. 5885798
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,868
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:
```

```

; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1184699
US-08-705-868-4

Query Match
Best Local Similarity 21.2%; Score 132.5; DB 2; Length 388;
Matches 79; Conservative 64; Mismatches 149; Indels 80; Gaps 19;

QY 60 HFLRRGI--FESHDMQVLDAYENKKPEFLDYTGSGPSSEAMHGHLPFI-FTKMLQDV 116
D 12 HLIFRNLOEVLGEERLKEILKERELK---IYWGATATGKR-HVAAYFVPMKIDFLKAG 66
QY 117 FNVPLVT---QMTDEKYLKDLTLDDQAYGDAVENAKDIIACGFDINKTFIFSDIDYMG 172
D 67 CEVTILFADLHAYLDNMKAPWELLELRVSYEENYIKAM-LESIGVLEKLFIFIGTDYQ- 124
QY 173 MSSGFKNVVVKIOKHVFNOYK---GIFGFTDSDICIKISFPALQAAFSFNSFPQIFR 228
D 125 LSKETTLDYRLSSVYVTHDSKKAGAEVYKQVEHPLISGLIYPLGIA---LDEEYLYK-- 179
QY 229 DRTIOCLIPCAIDDPYFRMTRDVAPRIGYKPPALHSTFFPALQAGQTKMSASDPNSS 288
D 180 ---DAQF---GIDQRKIFTFAEKYLPAIGYSKRVHLMNPMVPGITG--SKMSSSEESK 231
QY 289 IFLVDTAKOIKTKVNHAF-----SGGRDTIEH-----KQFGNCDVD 327
D 232 IDLDRKEDVKKLKK-AFCEPGVNNNGVLSFKHVLFPKSEFVYLDRKMGCN--- 286
QY 328 VSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALLIEVLOPLIAHQARRKEVTDEIV 387
D 287 ---KTYTAYVD-----LEKDFAAEVVHPGDLKNSVVALNKL-----DPIR 325
QY 388 KEFMTPR--RKLS 397
D 326 EKENTPALKILA 337

RESULT 7
US-09-123-615-4
; Sequence 4, Application US/09123615
; Patent No. 6090377
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,615
; FILING DATE:
```

US-09-492-581-2

Query Match	8.88;	Score 185.5;	DB 4;	Length 341;
Best Local Similarity	24.08;	Pred. No. 2.8e-12;		
Matches	81;	Conservative 60;	Mismatches 128;	Indels 69;
				Gaps 15;

QY	84	KPELYTJRGSSSEMHHGHLIPIETFWMLDVENVPLVIMTMDDEKY-LMKDITLLOQAY	142
Db	3	KPIIL-IGDRRTGK-LHIGHVGSLLKN-----VLQJEDKXYDMFVLADQAL	49
QY	143	GDAVEN-----AKDIACGFIDINKTIF--SDLDYMGSSGFKNV---KIQK	186
Db	50	TDHAKDPOTIVESIGNVALDLVAGLDPNKSTIISOIDPELALSMYMYNLVSLARLE	109
QY	187	HATENOVGIGFTDSDICIGISEPAAIOA--PSFSNSFPDIFRDRTDICIICARDOD	244
Db	110	NPLYKTEISQGFSGISPIPTGFLVYPIAQAADITAFKANY-----VPVGTDK	156
QY	245	PYFRKTRD-----VAPRICTPKPALHSTFPALOGAOTKMSASDPNSSIF	290
Db	157	PMIBOTREIVASFNNAYNCVDLVAPEGIYEPENE--RAGRLEGLDGG-NAKMSKS-LNNGIY	212
QY	291	LTDIAKQKTEYVKNHAFSGGRDITIEEHKOFGCNCDDVVSFMYLTFF--LEDDKLEQIRK	348
Db	213	LADADDLTRKKVVMGYMDPDHIREDDGKIEGN---MVFIYDLVFRPDADQADIAOMKE	266
QY	349	DYTSGAMLTGELKALIEVLOPLAEHQARKEVTDDEL	386
Db	269	RYORGGIGDVKTRKRLLEITREIRGLPIRREIRIEFAKOD	306

RESULT 4

US-08-743-130A-39
; Sequence 39, Application US/08743130A

```

1  GENERAL INFORMATION:
2  APPLICANT: Sassanfar, Mandana
3  APPLICANT: Gallant, Paul L.
4  APPLICANT: Shen, Xiaoyu
5  APPLICANT: Tao, Nianjun
6  APPLICANT: Tao, Jianshi
7  APPLICANT: Houman, Fariba
8  TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE
9  TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
10 NUMBER OF SEQUENCES: 41
11 CORRESPONDENCE ADDRESS:

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30A
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/743,130A

```

US-08-743-130A-39

Query Match	7.88;	Score 164.5;	DB 2;	Length 409;
Best Local Similarity	20.78;	Pred. No. 8.8e-10;		
Matches 78;	Conservative 59;	Mismatches 132;	Indels 107;	Gaps 16

```

Oy 76 VLDAEENK-KPEVLYATGSGPSEEMAHVGHILPFI-----FTKMLDQV-----F 117b
Db 27 IKDULEKENRPFVKIYWGTAFTPGKR-HCGYEPVMTKLAHFLKAGCEVYVLLADLHAFLADNM 85
Oy 118 NVPLVITQMTDDEKYLMLKDLTLDOAQYGADEVANAKDILACGEDINKTFEFLSDLYGMSGCF 177b
Db 86 KAPLEVVYVYRAKYEPFYVKKALIKSINVPIDELKREYVSSQYKGGDYV---MDLFLEKNIV 142b
Oy 178 YKNNVYK-----TQKHVTFNOVKKGIFFGLTSDSCICKISFPALQAAPEFSNSGFPQJEDRPTD 233b
Db 143 SQNDAKFRAGADVNOVAPNLLSGELT-----YPLMQA-----IDEHHLG 180b
Oy 233 IOCLIPCAIDDDPFERMRDVARPIRGYKRPALLHSTFFPALOGQOTKMSADPMSIETL 292b
Db 181 VDAQFG-GVDDRKJFEVLAEENLPSITGYKRAHLNMPVPGI-GGGKMSASDPMSKIDII 238b
Oy 293 DTAOKITKTVKHAHAFSG--RDT-----IEHRPGGN 323b
Db 239 EEPKYVKKKVASATCARBELKDNGLIAFIEVYIOPIAELKTGVGAFKDLIDREKKGG- 297b
Oy 324 CDVDVSFMYLTFLEDDDKLEIRKDYTGMLTSGMLTGLKALI---EVLQPLAENQARR 379b
Db 298 ---PLSY-----DSIQLKADFVDGKTLAPPDILKSGVADKINELLAPIRAEFESS- 343b
Oy 380 KEYTDEIKFEMTPRK 395
Db 344 -----EEFOVAK 351

```

RESULT 5

US-08-743-130A-2
; Sequence 2, Application US/08743130A

```

1  GENERAL INFORMATION:
2  APPLICANT: Sassanfar, Mandana
3  APPLICANT: Gallant, Paul L.
4  APPLICANT: Shen, Xiaoyu
5  APPLICANT: Tao, Nianjun
6  APPLICANT: Tao, Jianshi
7  APPLICANT: Houman, Fariba
8  TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE
9  TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
10 NUMBER OF SEQUENCES: 41
11 CORRESPONDENCE ADDRESS:
12

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.30
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/7743,130A
8  FILING DATE: 01-NOV-1996

```

```

; INFORMATION FOR SEQ ID NO: 39:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 409 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS:
;

```

Qy	64	RGIFESHMDMOVLDAEANKKPFELTYLRGPSSPAMJHGLIFETFKYLDVFNPLVY	123
Db	73	RGVFESQDRDLRLDLYEHGHPFLYLRGPSSDMLHGMVFETFKWLYEYFQVPIVY	132
Qy	124	QMTDEKYLKM-DLTDQANGDAVENAKDIIACGFDINTKFTITSDDLDMGMSGGFKRNV	182
Db	133	ELTDEKFLFYHQHJLTLDVYGFAEANKKDIIAVGFNEEMFTITSDLQYM--GAFENYV	190
Qy	183	KIOKHVFNQVKGJEGFTPDSCKIKISFPALQAAFSFNSFPQIFDRPTIOCLICAD	242
Db	191	RTSQITTSYAKAVFGLTDDSCIKIHFAISQIATNAPSSFPVGLGPRTPCLICAD	250
Qy	243	QDPFFKRTDRAPIGTGPKPALHSTFFPALQATMSADPNSSIFLDTAKQIKTY	302
Db	251	QDPFRCRDVADLRLFTKPKALHAKFFPALQGSTMSASDPTTSTFMDTKAQOIKTI	310
Qy	303	NKAHFSGGRIDEEHRQFGNCNDVAFNMLTFLEDDDLLEOIRKDYTSGAULTGELK	362
Db	311	NKYAFSGGRATAEEHRELGCNPEVDVAFQYLSFYSYDEKTLAQDEQYRGKILSGMK	370
Qy	363	ALIEVLOPLAEHQARKREYVDELYKRFMPRKISF	398
Db	371	ECITVLEEVSAVOERSKVDQDQVEKFMKHLV	406

RESULT 2
US-08-928-100-2
; Sequence 2, Application US/08928100

GENERAL INFORMATION:
APPLICANT: Gentry, Danlle
APPLICANT: Greenwood, Claire
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: Title
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,100
FILING DATE: 12-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9619072.3
FILING DATE: 12-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31624-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 341 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-100-2

Query Match	8.8%;	Score 185.5;	DB 3;	length 341;
Best Local Similarity	24.0%;	Pred. No. 2.8e-12;		

	Matches	81;	Conservative	60;	Mismatches	128;	Indels	69;	Gaps	15;
QY	84	KPEFLYTGRGSSSEAMHGHILFIETKWLQDVFNVPLVIOQTDEKY-LMKDLTLDOAY	142							
Dd	3	KPILL-TGDRLPTGK-LIHGHVGLSKNR-----VLLOEEDKDYMFVFLDAQOL	49							
QY	143	GDVAVN-----AKDIAGCFDINKFTIF--SDLDYGMSGGFKANY--KIQK	186							
Dd	50	TDHADDPQTIVESIGNVALDYLAVALDPKKSITFIQSQIPELIAELSYMYMNLVSALARL	109							
QY	187	HATFNOVGIFEGFTDDCIGKISPAIDAA--PSFSNSPFQIFRRDRTOICLICPAIDOD	244							
Dd	110	NPTVTEITSQGKGFGSITPGFLVYLIAQAADLTARKANT-----VPVGTIDOK	156							
QY	245	PYFRMTRD-----VADRIGYPKDALLSHFFPALOGAQTKMASDPNSSIT	290							
Dd	157	PMIEOTREIVRSFNNAVMCDVLEPEGIYPEHE--RAGRLPGLDG-NAKMSKS-LNNGY	212							
QY	291	LTDTAKKQIKTVKNKAIFSGGRTIDEHHQFGNCVDVDSFMULTFF--LEDDDKLEQLRK	348							
Dd	213	LADDSTLRKKWMSMYTDPDLIRVDDPKITEEN---WVFHYLLDVGEPRDAQOELADMKE	268							
QY	349	DYTSGAMLTELKALIEVLOPLIAEHQARREVDEMI	386							
Dd	269	RYORGGLGVCKTKRYLLEIRELERLPPIRERRRIEFPAKDM	306							

RESULT 3

; Sequence 2, Application US/09492581

```

1 GENERAL INFORMATION:
2 APPLICANT: Gentry, Danlie
3 APPLICANT: Greenwood, Claire
4 APPLICANT: Lawlor, Elizabeth
5 TITLE OF INVENTION: No. 6346409e1 tips
6 NUMBER OF SEQUENCES: 6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Smithkline Beecham Corporation
9 STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FastSeq for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/492,581
21 FILING DATE:
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US/08/928,100
25 FILING DATE: 12-SEP-1997
26 APPLICATION NUMBER: 96199072.3
27 FILING DATE: 12-SEP-1996
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Gimm1, Edward R
30 REGISTRATION NUMBER: 38,891
31 REFERENCE/DOCKET NUMBER: P31624-1
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 610-270-4478
34 TELEFAX: 610-270-5090
35 TELEX:
36 INFORMATION FOR SEQ ID NO: 2:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 341 amino acids
39 TYPE: amino acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 : Search time 9.8029 Seconds
(without alignments)
999.159 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471

Perfect score: 2116
Sequence: 1 SNHGPDATAEDEFEVDPWTV.....VTDEIVKFEFMPKLSDFDQ 401

Scoring table: BLASTUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCtUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1218.5	57.6	424	4	US-08-876-885-26
2	185.5	8.8	341	3	US-08-928-100-2
3	185.5	8.8	341	4	US-09-492-581-2
4	164.5	7.8	409	2	US-08-743-130A-39
5	162.5	7.7	409	2	US-08-743-130A-2
6	132.5	6.3	388	2	US-08-705-868-4
7	132.5	6.3	388	3	US-09-123-615-4
8	113.5	5.4	418	4	US-08-855-910-11
9	109	5.2	377	4	US-09-352-990-28
10	108	5.1	397	4	US-08-923-867-2
11	108	5.1	197	3	US-08-928-100-4
12	108	5.1	197	4	US-09-183-134-2
13	108	5.1	197	4	US-09-492-581-4
14	104.5	4.9	370	2	US-08-415-593-45
15	95	4.5	418	4	US-08-844-054-2
16	95	4.5	418	4	US-09-347-333-2
17	94.5	4.5	877	2	US-08-907-166-8
18	89.5	4.2	344	4	US-09-393-554-2
19	88.5	4.2	898	1	US-08-465-995A-4
20	88.5	4.2	898	2	US-08-465-994C-4
21	88.5	4.2	898	2	US-08-966-145-4
22	88.5	4.2	920	1	US-08-101-593-4
23	88	4.2	606	2	US-08-883-534-3
24	88	4.2	606	3	US-09-204-764-3
25	87.5	4.1	855	4	US-08-890-865A-10
26	87.5	4.1	898	1	US-08-465-995A-2
27	87.5	4.1	898	2	US-08-465-994C-2

28	87.5	4.1	898	2	US-08-966-145-2	Sequence 2, Appl1
29	87.5	4.1	920	1	US-08-101-593-2	Sequence 2, Appl1
30	85	4.0	410	4	US-09-352-990-16	Sequence 16, Appl1
31	83.5	3.9	428	3	US-08-331-625A-43	Sequence 43, Appl1
32	83.5	3.9	510	1	US-08-249-112-3	Sequence 3, Appl1
33	83.5	3.9	510	5	PCT-US95-06556-3	Sequence 7, Appl1
34	83.5	3.9	970	1	US-08-375-709-7	Sequence 7, Appl1
35	83.5	3.9	970	1	US-08-752-929-7	Sequence 7, Appl1
36	83.5	3.9	970	4	US-09-090-793-5	Sequence 2, Appl1
37	83.5	3.9	1452	4	US-08-331-625A-2	Sequence 2, Appl1
38	83.5	3.9	1452	5	PCT-US93-04384-18	Sequence 18, Appl1
39	83.5	3.9	1452	5	PCT-US93-04692-2	Sequence 18, Appl1
40	83	3.9	400	4	US-09-352-990-18	Sequence 18, Appl1
41	83	3.9	551	1	US-08-484-493-15	Sequence 15, Appl1
42	83	3.9	551	1	US-08-484-494-15	Sequence 15, Appl1
43	83	3.9	551	2	US-08-345-212-15	Sequence 15, Appl1
44	83	3.9	551	4	US-09-249-003-15	Sequence 15, Appl1
45	83	3.9	1073	1	US-07-695-564-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-876-885-26
: Sequence 26, Application US/08876885

: Patent No. 6174713

: GENERAL INFORMATION:

: APPLICANT: Shen, Xiaoyu

: APPLICANT: Houman, Fatima

: TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-TRNA

: TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING

: NUMBER OF INVENTIONS: SAME

: NUMBER OF SEQUENCES: 26

: CORRESPONDENCE ADDRESS:

: ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

: STREET: Two Militia Drive

: CITY: Lexington

: STATE: MA

: COUNTRY: USA

: ZIP: 02173

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patentin Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/876, 885

: FILING DATE: 16-JUN-1997

: CLASSIFICATION: 435

: ATTORNEY/AGENT INFORMATION:

: NAME: Brook, David E.

: REGISTRATION NUMBER: 22,592

: REFERENCE/DOCKET NUMBER: CP197-02

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (781) 861-6240

: TELEFAX: (781) 861-9540

: INFORMATION FOR SEQ ID NO: 26:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 424 amino acids

: TYPE: amino acid

: STRANDEDNESS:

: TOPOLOGY: linear

: US-08-876-885-26

Query Match 57.6%, Score 1218.5; DB 4; Length 424;
Best Local Similarity 57.1%, Pred. No. 3.4e-128;
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 8 TEAEEDFVDPWTVC---TSSAKGIDYDKLIVREGSSKIDKELNIRIRATGQRPHFRL 63
DB 13 TEESQKITPWEVGAVVDGKSMGIDYDKLISQFCKITTEBTLERFRQVTEGEPHPLK 72

Db 239 KEYGGRLTGEVKKRLTEVLTEIVEKHRRRAAAVTDEWDAFMAVRPLPSKFE 292

RESULT 13

AAB58517
ID AAB58517 standard; Protein; 173 AA.

AC AAB58517;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 855.

Human: lung cancer associated protein; neuroprotective; cytostatic;
cardioactive; immunomodulatory; muscular active; vulnerary;
gastrointestinal; nephrotoxic; antineoplastic; gynecological;
antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN W0200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.

PI Ruben SM;

XX WPI: 2000-587514/55.

DR N-PSDB; AAF18393.

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -

PS Claim 11; Page 1392-1393; 1425p; English.

CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58518. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotoxic; antineoplastic; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.

XX Sequence 173 AA;

XX Query Match 39.3%; Score 831; DB 21; Length 173;

XX Best Local Similarity 98.7%; Pred. No. 1.9e-77;

XX Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 107 FFFTKMLQDVFNPLVIMQTDKELKMDLTLDQAYGDAVENAKDIICGPDINKTFPS 166
Db 1 FFFTKMLQDVFNPLVIMQTDKELKMDLTLDQAYGDAVENAKDIICGPDINKTFPS 60

Qy 167 DLDYMGSSGFYKNNVYKIQKHVTFNQVKGIFGFTDSDICIGISPFAIQAAFSFNSFPQI 226
Db 61 DLDYMGSSGFYKNNVYKIQKHVTFNQVKGIFGFTDSDICIGISPFAIQAAFSFNSFPQI 120

Qy 227 FRDRTDIOCLIPCAIDODPYFRMTDVAVRIGYPPAL 264

Db 121 FRDRTDIOCLIPCAIDODPYFRMTDVAVRIGYPPAL 158

RESULT 14

AAB96409
ID AAB96409 standard; Protein; 385 AA.

AC AAB96409;

DT 29-OCT-2001 (first entry)

DE Putative P. abyssi tryptophanyl-tRNA synthetase.

DE Hyperthermophilic archaeon; hyperthermophilic protein.

XX Pyrococcus abyssi.

OS Pyrococcus abyssi.

PN FR2792651-A1.

PD 27-OCT-2000.

PF 21-APR-1999; 99FR-0005034.

PR 21-APR-1999; 99FR-0005034.

PA (CNRS) CNRS CENT NAT RECH SCI.
(IFRE-) IFREMER INST FR RECH EXPL MER.

PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Heilig R;

DR WPI: 2001-126236/14.

PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -

PS Claim 7; Pages 1099-1100; 1657p; French.

CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF66431 and AAF41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as W0200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAF75903-AAH75920 and AAG66436.

XX Sequence 385 AA;

XX Query Match 37.9%; Score 803; DB 22; Length 385;

XX Best Local Similarity 45.3%; Pred. No. 4.9e-74;

XX Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

Qy 12 EDF-VDPMVTQTSAGKIDYDKLIIVFGSSKIDKLIRIRIEATGGRHFLRNGIFPSH 70
Db 3 EDFKVTPEWEGV---VDYKNLIRHFGTSPTEBELKELTALTSSELPDFRRKFFPSH 58

Qy 71 RDMANQVDAYENKRPFYLYTGRGSSSEAMVHGHLIPFTKMLQDVFNPLVIMQTDK 130

Db 59 RDKYVLDQYDEGRGFLYTGGRGSG-PMHIGHIIPFATKWLQKFCGVNLYIQITDEK 117

Qy 131 YLWMD-LFLDQAYGAVENAKDIICGPDINKTFEFDLDYMGSSGFYKNNVYKIQKHT 189
Db 118 FLKENTLFEDTKHAYENILDIIVAGFDPKDTFLEONSE---TKTEMAIRPIAKKIN 173

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 45.7%; Score 968; DB 21; Length 292;
Best Local Similarity 60.9%; Pred. No. 2.9e-91;
Matches 179; Conservative 51; Mismatches 62; Indels 2; Gaps 1;

Qy 108 IFTKWLQDVFNVPVLTIOGTDEKYLKMDLTDOAVGDAVENAKDITACGPIKNTFIRS 167
Db 1 MFTKLOEAFKVPVLTIOGLTDEKSTYKMLSVESQRLARENKAKDITACGFDVTKTIFISD 60
Qy 168 LDYMGMSGFYKNVYKIOKHVTFNOVKCIFGFTDSDCIGKISFPALQAPFSNSFPOLF 227
Db 61 FDYVC--GAFYKNVYKVCVTLNKAKCIFGSGSDPIAKLSFPVQAVPSPSPFLFLF 118
Qy 228 RDRDIOCLIPCAIDDDYEFKMTDVAVRICYPKALHSTFFPALQCAOTKMSADPNS 287
Db 119 PGKDLRCLIPCAIDDDYEFMTDVAVRICYSKPALIESTFFPALQENKMSADPNS 178
Qy 288 SIFLDTAKOIKTYKNKAFSGGRDTTEHROFGNCDVDVSFWYLFELDDDKLEDR 347
Db 179 AIYVDSAKDIDKINRYAFSGGDSIEKHRELGNLEVDIPVKLSFLEDDSELEHRIK 238
Qy 348 KDYTSGAMLTELKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPKLSFDQ 401

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145214.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 61.6%; Score 1304; DB 21; Length 426;
Best Local Similarity 60.3%; Pred. No. 1e-125;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

QY 6 DATEAE--EDFVDPMVTQTSASAKIDYDKLIVREGSSKIDKELNRIERATGORPHHFLR 63
DB 31 DERAESEEOVYNPWEVSASAKDGKIDYDKLIDKFGCCORLDSLIDRVORLRSRPHVFLR 90
QY 64 RGIFFSHRDMNOVDAYEANKKFFLYLTGRGSSAMHGHILPFIETKWLQDVENVPLVI 123
DB 91 RSVFFAHRDFEILDAYERGRDKFFLYTGRGSSALHGHILPMPFKYIQEAKFVPLVI 150
QY 124 QMTDEKYLKDLTLDOAYGDAVENAKDIACGFDINKTFIIFSLDIPMGSSGFFYKNVYK 183
DB 151 QLTDEKSIWKNLSVESORLARENAKDIACGFDVTKTFIIFSDPVG--GAFYKKNVYK 208
QY 184 IOKHVTENOVKGIQGFDPDSCIGKISPPAIOAAPSFSNSFQIRDRTRDIOCLPCAIDQ 243
DB 209 VKKCVTLKAMGIRGFGSEDDIYAKLSPPVOAVSFFSSFFHLDPGKONLCILPCAIDQ 268
QY 244 DPFYFMTRDVAPRIGYKSPALHSTFFPALOGAOTKMSASDPNSSIFLTDIAKOIKTVN 303
DB 269 DPFYFMTRDVAPRIGYKSPALHSTFFPALOGENGKMSASDPNSAIVTSDAKDIKKIN 328
QY 304 KHAFSGRDTIEHRQFGNCDVVSFMYLTFFLEDDDKLEQIKDYTSGAMLTELKKA 363
DB 329 RYAFSGGDSIEKHELGANLEVDIPKYLSFFLEDSSELIHKEVGEGRMLTGEVKKR 388
QY 364 LIEVLOPLIAEHQARRKEVYDEIVKEFETPRKLSFDDQ 401
DB 389 LIEVLTIEVEKHRRARAAYIDEMVADAFMAVAPLPKFE 426
RESULT 11
ID AAB66931 standard; Protein; 424 AA.
XX AAB66931;
XX AC AAB66931;
XX

```
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 61.6%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 9,5e-126;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

QY 6 DATEAE--EDVVDPTVOTSSAKGIDYKLVREGSSKIDKELNIRATGQRPHEFLR 63
   | ||| | :||| | ||||| :|| :||:| | :|| |||
Db 7 DERAESESOYVNPWEVSADKDGKIDYKLDKFCQRLDLSLDYRQRLTSROPHEFLR 66
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
QY 64 RGIFSHDMNOVLDAFENKKPFYLYTGRGSPSEAMVGHILPIFTFKWLDVFNVLVI 123
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
Db 67 RSVFFAHDFNEILDAYRGDKFYLYTGRGSPSEALHGHILPFYTYLDEAFNVLVI 126
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
QY 124 QMTDEKYLKMDLFDQAYGDAVENAKDIACGFDINKFTFSDIDYGMSSGFYKNVYK 183
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
Db 127 QLTDEKSIWKRLSEESQRLARENAKDIIACGFVYTKTFISFSDYVG--GAFYKNVYK 184
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
QY 184 IOKHTFNOVNGIGFETDSDIGKISFPALQAPSENSFOIEFDRDIOCLIPCAIDQ 243
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
Db 185 VGKCVTLTKAMGIFGFSGEDPIAKISFPPOAVSFSPSFPFLFGKONLRCLIPCAIDQ 244
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
QY 244 DPFYMRDVAAPRIGYPRPALHSTFPALOGAQTMSASDPNSSIFLDTAKQIKTRVN 303
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
Db 245 DPFYMRDVAAPRIGYSKALIESTFPPALOGENGKMSASDPNSAIYVTSKAKDINKIN 304
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
QY 304 KHAFSGGNDTIEHRQFGNCDVVSFMYLTFLEDDDKLEQIRKDYSGAMLTGELKKA 363
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
Db 305 RYAFSSGGDSIEKHRELGANLEVIDIPVKYLSFLEDDDELHIKKEYGEGRLTGEVKKR 364
   | :||| | :||:| | :||| | ||||| ||||| :|||:| | :||| |
QY 364 LIEVLOPLIAEHOARKREVTDEIYKEFMTPKRISFDQ 401
   | ||| :| :| :| | ||||| | | | | |
Db 365 LIEVLTETVEKRRARAATDEMDAFMAVRPLPSKFE 402
   | ||| :| :| :| | ||||| | | | | |

RESULT 10
AAG23697
ID AAG23697 standard; Protein: 426 AA.
XX
AC AAG23697;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SMO ID NO: 27100.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
XX
```

```
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126254.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 08-APR-1999; 990S-0128234.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 18-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134376.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139753.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR
```

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132488.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

KW vascular endothelial cell function; burn; plastic surgery; abdomen;
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
 KW dermal ulcer; diabetic ulcer; endothelialization;
 KW lypophanyl-tRNA synthetase; tRPS; vascular graft surgery.
 XX
 OS Homo sapiens.
 PN WO200174841-A1.
 XX
 PD 11-OCT-2001.
 XX
 XX 21-MAR-2001; 2001WO-US08966.
 XX
 XX 31-MAR-2000; 2000US-193471P.
 XX
 XX (SCRI) SCRIPPS RES INST.
 XX
 PI Schimmel P, Wakasugi K;
 XX
 XX WPI: 2001-626377/72.
 DR N-PSDB: AAB43605.
 XX
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
 PT vascular endothelial function, in particular for regulating
 PT angiogenesis, tumor metastasis and treating myocardial infarction -
 XX
 XX Disclosure: Page 135-36; 150pp; English.
 PS
 XX The sequences given in AAB47615-18 show full length and truncated
 CC versions of lypophanyl-tRNA synthetase (tRPS). The truncated tRPS of
 CC the invention comprises a Rossmann fold nucleotide binding domain, and
 CC is capable of regulating vascular endothelial cell function. It is of
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
 CC full length tRPS with polymorphonuclear leucocyte elastase. Truncated
 CC tRPS is useful for regulating angiogenesis, tumor metastasis,
 CC enhancing angiogenesis to a graft, treating myocardial infarction,
 CC solid tumor, and a condition that would benefit from increased or
 CC decreased angiogenesis in a mammal, in particular humans. It is also
 CC useful in diagnosis and as a wound healing agent for treating wounds
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
 CC plastic surgery when reconstruction is required following a burn or
 CC for cosmetic purposes. It is particularly useful in the treatment of
 CC abdominal wounds where there is high risk of infection. Truncated tRPS
 CC promotes endothelialization in vascular graft surgery and is used in
 CC conjunction with angiography to administer the angiogenic tRNA
 CC synthetase polypeptides or polynucleotides directly to the lumen and
 CC wall of the blood vessel.
 CC
 XX
 XX Sequence 392 AA;
 SQ
 Query Match 94.0%; Score 1988; DB 22; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2,1e-196;
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 24 SAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFLRGIFSSHDMQVDAENK 83
 DB 2 SAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFLRGIFSSHDMQVDAENK 61
 QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVFNVPVLIOMTDEKYLMDLTLDQAYG 143
 DB 62 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVFNVPVLIOMTDEKYLMDLTLDQAYG 121
 QY 144 DAVENAKDIACGFDINKTFFSDDLIDYGMSSGFYKNNVKIQKHVTENQVKGIFGFTDSD 203
 DB 122 DAVENAKDIACGFDINKTFFSDDLIDYGMSSGFYKNNVKIQKHVTENQVKGIFGFTDSD 181
 QY 204 CIGKISPAIOAAPSFSFPOIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPPA 263
 DB 182 CIGKISPAIOAAPSFSFPOIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPPA 241
 QY 264 LLSHTFFPALOGAOTKMSADPNSSIFLTDPAKOIKTKVNHAFSGGRTIEEHROFCGN 323
 DB 264 LLSHTFFPALOGAOTKMSADPNSSIFLTDPAKOIKTKVNHAFSGGRTIEEHROFCGN 323

DB 242 LLSHTFFPALOGAOTKMSADPNSSIFLTDPAKOIKTKVNHAFSGGRTIEEHROFCGN 301
 QY 324 CDVDSFMYLTFLEDDDKLEOIKRDYSGAMLTGELKKALIEVLOPIAEHOARRKEVT 383
 DB 302 CDVDSFMYLTFLEDDDKLEOIKRDYSGAMLTGELKKALIEVLOPIAEHOARRKEVT 361
 QY 384 DEIVKEFMTPRKLSPDFQ 401
 DB 362 DEIVKEFMTPRKLSPDFQ 379
 RESULT 7
 ABB64621
 ID ABB64621 standard; Protein; 430 AA.
 XX
 AC ABB64621;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 20655.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 OS
 XX Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 DR N-PSDB: ABL08724.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Disclosure: SEQ ID NO 20655; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 XX Sequence 430 AA;
 SQ
 Query Match 64.8%; Score 1370.5; DB 22; Length 430;
 Best Local Similarity 64.3%; Pred. No. 1.4e-132;
 Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;
 QY 7 ATAEEDPVDPMVOTSSAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFLRGIGI 66
 DB 38 ATAEEDPVDPMVOTSSAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFLRGIGI 97
 QY 67 FFSHRDMNQVDAENKRPYLYTGRGSSSEAMHVGHLIPFTFKWLQDVFNVPVLIOMT 126
 DB 98 FFSHRDLHTLTLTREQCKPFLYLYTGRGSSSEAMHVGHLIPFTFKWLQDVFNVPVLIOMT 157

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytoskeletal; cardioprotective;
CC immunomodulatory; muscular active general; vulnery; gastrointestinal
CC general; nephrotropic; anti-infective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 475 AA:
Query Match 99.3%; Score 2101; DB 21; Length 475;
Best Local Similarity 99.5%; Pred. No. 5.9e-208;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPH 60
DB 75 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPH 134
QY 61 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 120
DB 135 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 194
QY 121 LVIOGTDDEKYLKMDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 135 LVIOGTDDEKYLKMDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 254
QY 181 VVKIQKHTFNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFDRDIDQCLIPCA 240
DB 255 VVKIQKHTFNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFDRDIDQCLIPCA 314
QY 241 IDDDYFPMTRDVAPRIGPKPALLHSTFFPALQAGQTKMSADPNSSIFLDTAKQIKT 300
DB 315 IDDDYFPMTRDVAPRIGPKPALLHSTFFPALQAGQTKMSADPNSSIFLDTAKQIKT 374
QY 301 KVNKAHFGSGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEDIRDYTSGAMLTGEL 360
DB 375 KVNKAHFGSGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEDIRDYTSGAMLTGEL 434
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSDFQ 401
DB 435 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSDFQ 475
RESULT 5
AAY05372
ID AAY05372 standard; Protein: 471 AA.
XX
AC AAY05372;
XX
DT 30-JUN-1999 (first entry)
XX
DE Human HCMV inducible gene protein, SEQ ID NO 12.
XX
KW HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;
XX anti-viral therapy; anti-HCMV therapy; detection; diagnosis;
XX drug screening.
XX
OS Homo sapiens.
XX
XX WO9913075-A2.
XX
XX PD 18-MAR-1999.
XX

PF 08-SEP-1998; 98WO-US18638.
XX
PR 22-SEP-1997; 97US-0059725.
PR 08-SEP-1997; 97US-0058180.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Cong J, Schenk T, Zhu H;
XX
XX WPI: 1999-243729/20.
DR N-PSDB; AAX33942.
XX
PT New isolated human genes
PS Claim 3; Page 112-114; 184pp; English.
XX
CC This sequence is encoded by a human gene of the invention, and is induced
CC to express by both HCMV and Interferon (IFN), designated HCMV-Inducible
CC genes (c1g or c1gs). The invention also relates to genes that are
CC repressed in the presence of HCMV infection, designated HCMV-repressible
CC genes (c1rg or c1rgs). The products can be used to obtain agents which can
CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can
CC also be used for the development of drugs that would allow for higher
CC dosage IFN treatments without the concomitant toxicity normally
CC associated with administering high levels of IFN. The products can also
CC be used for detection, diagnosis and drug screening.
XX
SQ Sequence 471 AA:
Query Match 99.1%; Score 2096; DB 20; Length 471;
Best Local Similarity 99.3%; Pred. No. 1.9e-207;
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPH 60
DB 71 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPH 130
QY 61 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 120
DB 131 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 190
QY 121 LVIOGTDDEKYLKMDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 191 LVIOGTDDEKYLKMDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
QY 181 VVKIQKHTFNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFDRDIDQCLIPCA 240
DB 251 VVKIQKHTFNOVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFDRDIDQCLIPCA 310
QY 241 IDDDYFPMTRDVAPRIGPKPALLHSTFFPALQAGQTKMSADPNSSIFLDTAKQIKT 300
DB 311 IDDDYFPMTRDVAPRIGPKPALLHSTFFPALQAGQTKMSADPNSSIFLDTAKQIKT 370
QY 301 KVNKAHFGSGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEDIRDYTSGAMLTGEL 360
DB 371 KVNKAHFGSGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEDIRDYTSGAMLTGEL 430
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSDFQ 401
DB 431 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSDFQ 471
RESULT 6
AAB47618
ID AAB47618 standard; Protein: 392 AA.
XX
AC AAB47618;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human inactive TTPRS.
XX
KW Tyrosyl-tRNA synthetase; TTPRS; Rossmann fold nucleotide binding domain;

QY 301 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEQIRKDYTSGAMLTGEL 360
DB 324 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEQIRKDYTSGAMLTGEL 383
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401
DB 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 424

RESULT 3
AAB47615
ID AAB47615 standard; Protein: 484 AA.
AC AAB47615;
XX 07-JAN-2002 (first entry)
DE Human full-length TrpRS.
XX Tyrosyl-tRNA synthetase: TrpRS: Rossmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.
XX Homo sapiens.
OS
XX MO200174841-A1.
PN 11-OCT-2001.
PD 21-MAR-2001; 2001WO-US08966.
XX 31-MAR-2000; 2000US-193471P.
PR (SCRI) SCRIPPS RES INST.
XX (SCRI) SCRIPPS RES INST.
PI Schimmel P, Wakasugi K;
XX WPI: 2001-626377/72.
DR N-PSDB: AAH43602.
XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
XX
XX Disclosure: Page 117-19; 150pp; English.

CC The sequences given in AAB47615-18 show full length and truncated
CC versions of triphophanyl-tRNA synthetase (trpRS). The truncated trpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kDa molecular weight and is produced by cleavage of
CC full length trpRS with polymorphonuclear leucocyte elastase. Truncated
CC trpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated trpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.

XX Sequence 484 AA;
SQ Query Match 100.0%; Score 2116; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 1,7e-209;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNHGDATEAEEDPVDPPTVOTSSAKGIDYDKLYRVSSSKIDKELINRIERATGQRPH 60
DB 71 SNHGDATEAEEDPVDPPTVOTSSAKGIDYDKLYRVSSSKIDKELINRIERATGQRPH 130
QY 61 FLRRCIFFSHRDMNOVLDAVENKKRFYLYTGRGPSSSEAMHGHLIPFTFKMLQDFVNP 120
DB 131 FLRRCIFFSHRDMNOVLDAVENKKRFYLYTGRGPSSSEAMHGHLIPFTFKMLQDFVNP 190
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKTFTFSLDYMGSSGFYKN 180
DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKTFTFSLDYMGSSGFYKN 250
QY 181 VKIKQHVTFNOVKCIFGFTSDICGKISFPAIQAPFSNSFQIFRDRDIQCLPCA 240
DB 251 VKIKQHVTFNOVKCIFGFTSDICGKISFPAIQAPFSNSFQIFRDRDIQCLPCA 310
QY 241 IDODEYFRMTDVAAPRIGYPPALHSTFPALQAGOTKMSASDPNSSIFLTPRAKOIKT 300
DB 311 IDODEYFRMTDVAAPRIGYPPALHSTFPALQAGOTKMSASDPNSSIFLTPRAKOIKT 370
QY 301 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEQIRKDYTSGAMLTGEL 360
DB 371 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEQIRKDYTSGAMLTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401
DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 471

RESULT 4
AAB58220
ID AAB58220 standard; Protein: 475 AA.
AC AAB58220;
XX 14-MAR-2001 (first entry)
DE Lung cancer associated polypeptide sequence SEQ ID 558.
XX Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardiactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.
XX Homo sapiens.
OS
XX WO200055180-A2.
PN 21-SEP-2000.
PD 08-MAR-2000; 2000WO-US05918.
XX 12-MAR-1999; 99US-0124270.
PR (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
PA Ruben SM;
XX WPI: 2000-587514/55.
DR N-PSDB: AAF18096.
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX Claim 11; Page 1052-1053; 1425pp; English.
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
PS
XX Disclosure; Page 129-30; 150pp; English.
CC The sequences given in ABA47615-18 show full length and truncated
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.
XX
SQ Sequence 415 AA;
Query Match 100.0%; Score 2116; DB 22; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.4e-209;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNHGPDATAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRIRATGORPHH 60
DB 2 SNHGPDATAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRIRATGORPHH 61
QY 61 FLRRGIFFSHRDMNOVDLAYENKKPFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 120
DB 62 FLRRGIFFSHRDMNOVDLAYENKKPFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 121
QY 121 LVIOHTDEKYLKMDLTLDQAYGDAVENAKDIACGPIINTFTFSDIDYGMSSGFYKN 180
DB 122 LVIOHTDEKYLKMDLTLDQAYGDAVENAKDIACGPIINTFTFSDIDYGMSSGFYKN 181
QY 181 VKIÖKHVTFNQVKGIFGFTSDICGKISPAIOAAPSFSNFPQIFRDRDIOCLIPCA 240
DB 182 VKIÖKHVTFNQVKGIFGFTSDICGKISPAIOAAPSFSNFPQIFRDRDIOCLIPCA 241
QY 241 IDDPYFRMTDVAAPRIGYPRPALHSTFFPALOGAOTKMSASDPNSSIFLTDRAKÖIKT 300
DB 242 IDDPYFRMTDVAAPRIGYPRPALHSTFFPALOGAOTKMSASDPNSSIFLTDRAKÖIKT 301
QY 301 KVNKHAFFSGGDDTIEHHQFGNCDVDSFMYLTFLEDDDKLEDIRKDYSGAMLTGEL 360
DB 302 KVNKHAFFSGGDDTIEHHQFGNCDVDSFMYLTFLEDDDKLEDIRKDYSGAMLTGEL 361
QY 361 KKALIEVLÖPLIAEHQARKKEVTDEIVKEFMTPRKLSFDFO 401
DB 362 KKALIEVLÖPLIAEHQARKKEVTDEIVKEFMTPRKLSFDFO 402
RESULT 2
ID ABA47616 standard; Protein; 437 AA.
XX
AC ABA47616;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human mini TrpRS.
XX
KW Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;

KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW triphophanyl-tRNA synthetase; trpRS; vascular graft surgery.
XX
OS Homo sapiens.
XX
PN WO200174841-A1.
XX
PD 11-Oct-2001.
XX
PE 21-MAR-2001; 2001WO-US08966.
XX
PR 31-MAR-2000; 2000US-193471P.
XX
PA (SCRI) SCRIIPS RES INST.
XX
PI Schimmel P, Wakasugi K;
XX
DR WPI; 2001-626377/72.
DR N-PSDB; AAH43603.
XX
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction -
PS
XX Disclosure; Page 123-24; 150pp; English.
CC The sequences given in ABA47615-18 show full length and truncated
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Rossmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,
CC enhancing angiogenesis to a graft, treating myocardial infarction,
CC solid tumor, and a condition that would benefit from increased or
CC decreased angiogenesis in a mammal, in particular humans. It is also
CC useful in diagnosis and as a wound healing agent for treating wounds
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in
CC plastic surgery when reconstruction is required following a burn or
CC for cosmetic purposes. It is particularly useful in the treatment of
CC abdominal wounds where there is high risk of infection. Truncated TrpRS
CC promotes endothelialization in vascular graft surgery and is used in
CC conjunction with angiography to administer the angiogenic tRNA
CC synthetase polypeptides or polynucleotides directly to the lumen and
CC wall of the blood vessel.
XX
SQ Sequence 437 AA;
Query Match 100.0%; Score 2116; DB 22; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.5e-209;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNHGPDATAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRIRATGORPHH 60
DB 24 SNHGPDATAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRIRATGORPHH 83
QY 61 FLRRGIFFSHRDMNOVDLAYENKKPFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 120
DB 84 FLRRGIFFSHRDMNOVDLAYENKKPFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 143
QY 121 LVIOHTDEKYLKMDLTLDQAYGDAVENAKDIACGPIINTFTFSDIDYGMSSGFYKN 180
DB 144 LVIOHTDEKYLKMDLTLDQAYGDAVENAKDIACGPIINTFTFSDIDYGMSSGFYKN 203
QY 181 VKIÖKHVTFNQVKGIFGFTSDICGKISPAIOAAPSFSNFPQIFRDRDIOCLIPCA 240
DB 204 VKIÖKHVTFNQVKGIFGFTSDICGKISPAIOAAPSFSNFPQIFRDRDIOCLIPCA 263
QY 241 IDDPYFRMTDVAAPRIGYPRPALHSTFFPALOGAOTKMSASDPNSSIFLTDRAKÖIKT 300
DB 264 IDDPYFRMTDVAAPRIGYPRPALHSTFFPALOGAOTKMSASDPNSSIFLTDRAKÖIKT 323

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 25.1199 Seconds
(without alignments)
1773.117 Million cell updates/sec

Title: US-09-813-718-10_COPY_71_471
Perfect score: 2116
Sequence: 1 SNHGPDATAEEDFVDPMTV.....VTDEIVKEMTPKLSFDQ 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
A.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	415	22	AA847617 Human supermini tr
2	2116	100.0	437	22	AA847616 Human mini trPRS.
3	2116	100.0	484	22	AA847615 Human full-length
4	2101	99.3	475	21	AA858220 Lung cancer associ
5	2096	99.1	471	20	AAV05372 Human HCMV inducib
6	1988	94.0	392	22	AA847618 Human inactive Trp
7	1370.5	64.8	430	22	ABB64621 Drosophila melanog
8	1370.5	64.8	430	22	ABB67203 Drosophila melanog
9	1304	61.6	402	21	AA823698 Arabidopsis thalia
10	1304	61.6	426	21	AA823697 Arabidopsis thalia
11	1218.5	57.6	424	22	AA866931 Tryptophanyl-tRNA

12	968	45.7	292	21	AA823699 Arabidopsis thalia
13	831	39.3	173	21	AA858517 Lung cancer associ
14	803	37.9	385	22	AA896409 putative P. abysssi
15	226.5	10.7	142	21	AA858219 Lung cancer associ
16	185.5	8.8	341	19	AA856423 Tryptophanyl tRNA
17	185.5	8.8	341	22	AA837851 Streptococcus pneu
18	164.5	7.8	409	20	AA894248 C. albicans tyrosy
19	162.5	7.7	409	20	AA894247 C. albicans tyrosy
20	159.5	7.5	379	20	AA837623 protein involved i
21	158.5	7.5	379	22	AA896600 putative P. abysssi
22	152	7.2	344	20	AA835439 Chlamydia pneumoni
23	146	6.9	339	22	AA835829 Helicobacter pylor
24	138.5	6.5	140	21	AA823032 Arabidopsis thalia
25	134.5	6.4	525	22	AA860745 Drosophila melanog
26	134	6.3	372	22	AA847614 Human mini trPRS m
27	132.5	6.3	372	22	AA847612 Human mini trPRS.
28	132.5	6.3	356	22	AA847611 Human full length
29	128	6.0	334	22	AA835477 Haemophilus influe
30	123	5.8	347	21	AA831904 Arabidopsis thalia
31	122.5	5.8	346	21	AA848617 Arabidopsis thalia
32	118.5	5.6	404	22	AA854020 Propionibacterium
33	115.5	5.5	129	21	AA823033 Arabidopsis thalia
34	113.5	5.4	418	22	AA820578 Enterococcus tyros
35	113.5	5.4	423	22	AA835085 Enterococcus faeca
36	113	5.3	415	22	AA833998 Staphylococcus aur
37	113	5.3	420	18	AA819781 Tyrosyl-tRNA synth
38	113	5.3	420	22	AA837056 Staphylococcus aur
39	112	5.3	345	22	AA830498 C glutamicum prote
40	108	5.1	334	22	AA838213 Salmonella typhi c
41	106.5	5.0	370	22	AA834778 E. coli cellulular p
42	104.5	4.9	370	20	AA817508 Pneumocystis carin
43	100.5	4.7	421	22	AA823633 S. epidermidis ope
44	95	4.5	418	18	AA827663 Streptococcus pneu
45	95	4.5	418	22	AA837869 Streptococcus pneu

ALIGNMENTS

RESULT 1	AA847617 standard; Protein: 415 AA.
XX	AA847617:
XX	07-JAN-2002 (first entry)
XX	Human supermini trPRS.
DE	Tyrosyl-tRNA synthetase; TyRS; Rosemann fold nucleotide binding domain;
XX	vascular endothelial cell function; burn; plastic surgery; abdomen;
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW	dermal ulcer; diabetic ulcer; endothelialization;
KW	tryptophanyl-tRNA synthetase; trPRS; vascular graft surgery.
XX	
OS	Homo sapiens.
XX	
PN	WO200174841-A1.
XX	
PD	11-OCT-2001.
XX	
PF	21-MAR-2001; 2001WO-US08966.
XX	
PR	31-MAR-2000; 2000US-193471P.
XX	
PA	(SCRI) SCRIpps RES INST.
XX	
PI	Schimmel P, Wakasugi K;
XX	
DR	WPI: 2001-626377/72.
XX	
DR	N-PSDB: AA843604.
XX	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 14.2521 Seconds
(without alignments)
2858,658 Million cell updates/sec

Title: US-09-813-718-10_COPY_48_471

Perfect score: 2246
Sequence: 1 MSKRAAGEDTKADCPGPNP.....VTDEIVKEFMTPKLSFDQ 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR-71:*

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	99.3	471	1	AA1706
2	2134.5	95.0	475	1	YBMO
3	2041	90.9	475	1	YWRBPR
4	2032	90.5	481	2	S50053
5	1210	53.9	395	2	S58157
6	1163	51.8	432	2	S51901
7	907	40.4	386	2	C90190
8	803	35.8	385	2	C75020
9	626.5	27.9	301	2	G71206
10	534.5	23.8	380	2	G84373
11	409.5	18.2	370	2	PG4476
12	397.5	17.7	364	2	BE9131
13	386	17.2	134	2	T43806
14	370.5	16.5	420	2	E69461
15	353	15.7	374	2	D72477
16	274.5	12.2	513	2	F84371
17	192	8.5	323	2	H69346
18	188	8.4	364	2	E72512
19	185.5	8.3	341	2	D95250
20	185.5	8.3	341	2	G98125
21	174.5	7.8	341	2	B86633
22	172	7.7	394	2	A45999
23	170.5	7.6	366	2	S75410
24	169	7.5	395	2	H70385
25	167	7.4	351	2	E75438
26	166.5	7.4	346	2	B71496
27	164.5	7.3	895	2	A86410
28	163.5	7.3	353	2	E70100
29	163	7.3	460	2	C84750

30	159.5	7.1	337	2	F71300	tryptophan--trna 1
31	158.5	7.1	375	2	B75072	tyrosyl--trna synth
32	158	7.0	346	2	C81654	tryptophanyl--trna
33	154.5	6.9	102	2	T44994	probable tryptophan
34	152	6.8	344	2	H86590	tryptophanyl trna
35	152	6.8	344	2	C72034	tryptophan--trna 1
36	149.5	6.7	337	2	C84374	tryptophan--trna 1
37	149.5	6.7	408	2	T03741	tyrosyl--trna synth
38	146	6.5	339	2	E64676	probable tyrosine--
39	146	6.5	375	2	F71093	tryptophan--trna 1
40	145	6.5	337	2	A11066	tyrosine--trna 1
41	138	6.1	319	2	H69102	tryptophan--trna 1
42	135.5	6.0	328	2	C72370	tryptophan--trna 1
43	135	6.0	365	2	E82052	tryptophanyl--trna
44	132.5	5.9	343	2	S73024	tryptophan--trna 1
45	132	5.9	328	1	YMBSF	tryptophan--trna 1

ALIGNMENTS

RESULT 1
AA1706
tryptophan--trna ligase (EC 6.1.1.2) [similarity] - human
M:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor 1
C:Species: Homo sapiens (man)
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #ext_change 19-May-2000
C:Accession: A41633; AA1706; S19246; JN0676; JH0533; S26287
R:Flückner, J.; Rasmussen, H.H.; Justesen, J.
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991
A:Title: Human Interferon gamma potentially induces the synthesis of a 55-kDa protein (C
A:Reference number: A41633; M01D:92107982
A:Accession: A41633
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821
J:Rudolf, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.
J. Biol. Chem. 266, 24245-24248, 1991
A:Title: Interferon induces tryptophanyl--trna synthetase expression in human fibrobl
A:Reference number: A41706; M01D:92105071
A:Accession: A41706
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-471 <RUB>
A:Cross-references: GB:X67918; NID:932708; PIDN:AAA67324.1; PID:9184657
R:Buwitt, U.; Flohr, T.; Boettger, E.C.
EMBO J. 11, 489-496, 1992
A:Title: Molecular cloning and characterization of an interferon induced human cDNA v
A:Reference number: S19246; M01D:92164636
A:Accession: S19246
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-423, 'R', 'A25-471 <BOW>
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709
A:Note: 213-Ser and 214-Tyr were also found
R:Frolova, L.Y.; Grigorjeva, A.Y.; Sudomolna, M.A.; Kisselev, L.L.
Gene 188, 237-245, 1993
A:Title: The human gene encoding tryptophanyl--trna synthetase: Interferon-response e
A:Reference number: JN0676; M01D:93292932
A:Accession: JN0676
A:Molecule type: DNA
A:Residues: 1-141;182-471 <FROL>
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X6792
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB
A:Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residu
R:Frolova, L.Y.; Sudomolna, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kisselev, L.L.
Gene 109, 291-296, 1991
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr
A:Reference number: JH0533; M01D:92112058
A:Accession: JH0533
A:Molecule type: mRNA

A:Residues: 1-212, 'GD', 215-471 <FRO2>
 A:Cross-references: GB:M61715; NID:9340367; PIDN:AAA61298.1; PID:9340368
 A:Experimental source: fibroblast
 C:Genetics:
 A:Gene: GDB:MARS; IFP53
 A:Cross-references: GDB:119632; OMIM:191050
 A:Map position: 14q23-14q31
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;
 Best Local Similarity 99.5%; Pred. No. 146-173;
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MSYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRG 60
DB 48 MSYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRG 107
61 GSSKIDKELINRIERATGPRHFLRGIFFSHRDMNOVDAYENKKPFYLYTGRGSSA 120
DB 108 GSSKIDKELINRIERATGPRHFLRGIFFSHRDMNOVDAYENKKPFYLYTGRGSSA 167
OY 121 AMYGHILPIFTFKLQDVENVPLVYIOMTDEKYLKDLTDQAYGDAVENAKDIACGF 180
DB 168 AMYGHILPIFTFKLQDVENVPLVYIOMTDEKYLKDLTDQAYGDAVENAKDIACGF 227
OY 181 DINKTFISDLIDYMGSSGFYKKNVYKIOKHVTFNOVYKIGFTSDICIGISFPALQAP 240
DB 228 DINKTFISDLIDYMGSSGFYKKNVYKIOKHVTFNOVYKIGFTSDICIGISFPALQAP 287
OY 241 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRKPALHSTFPALQAG 300
DB 288 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRKPALHSTFPALQAG 347
OY 301 TKMSASDPNSSIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFEFL 360
DB 348 TKMSASDPNSSIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFEFL 407
OY 361 EDDDKLEQIRKDYTSAGMLTSELKALIEVLOPLIAEHQARRKEVTEIYKEEMTPRKLS 420
DB 408 EDDDKLEQIRKDYTSAGMLTSELKALIEVLOPLIAEHQARRKEVTEIYKEEMTPRKLS 467
OY 421 FDFQ 424
DB 468 FDFQ 471

```

RESULT 2

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - bovine
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 30-Sep-1992 #sequence, revision 30-Sep-1992 #text, change 21-Jul-2000
 C:Accession: A40279; JN0354; S10460; S14540
 R:Garret, M.; Pajot, B.; Trezeget, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti
 Biochemistry 30, 7809-7817, 1991
 A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic
 A:Reference number: A40279; MUID:91329348
 A:Accession: A40279
 A:Molecule type: mRNA
 A:Residues: 1-475 <GAR>
 A:Cross-references: GB:M74074; EMBL:X53918; NID:9163798; PIDN:AAA30799.1; PID:9163799
 A:Experimental source: pancreas
 A:Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue
 A:Note: part of this sequence was confirmed by protein sequencing
 R:Zargava, T.A.; Kovalova, G.K.; Pavlova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.
 Biorg. Khim. 15, 1307-1311, 1989
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca
 A:Reference number: JN0354; MUID:90211408
 A:Accession: JN0354
 A:Molecule type: protein

A:Residues: 112-124;282-287, 'N', 288, 'F', 289-292, 'Q', 293-294, 'IR', 336-353;423-441,443-
 A:Experimental source: liver
 A:Note: this paper is in Russian
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 F:24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;
 Best Local Similarity 95.0%; Pred. No. 146-165;
 Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

```

OY 2 SYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRG 61
DB 54 SYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYDKLYRG 113
62 SSKIDKELINRIERATGPRHFLRGIFFSHRDMNOVDAYENKKPFYLYTGRGSSA 121
DB 114 SSKIDKELINRIERATGPRHFLRGIFFSHRDMNOVDAYENKKPFYLYTGRGSSA 173
OY 122 MYGHILPIFTFKLQDVENVPLVYIOMTDEKYLKDLTDQAYGDAVENAKDIACGF 181
DB 174 MYGHILPIFTFKLQDVENVPLVYIOMTDEKYLKDLTDQAYGDAVENAKDIACGF 232
OY 182 INKTFISDLIDYMGSSGFYKKNVYKIOKHVTFNOVYKIGFTSDICIGISFPALQAP 241
DB 233 INKTFISDLIDYMGSSGFYKKNVYKIOKHVTFNOVYKIGFTSDICIGISFPALQAP 292
OY 242 FNSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRKPALHSTFPALQAG 301
DB 293 FNSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRKPALHSTFPALQAG 352
OY 302 KMSASDPNSSIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFEFL 361
DB 353 KMSASDPNSSIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFEFL 412
OY 362 DDDKLEQIRKDYTSAGMLTSELKALIEVLOPLIAEHQARRKEVTEIYKEEMTPRKLS 421
DB 413 DDDKLEQIRKDYTSAGMLTSELKALIEVLOPLIAEHQARRKEVTEIYKEEMTPRKLS 472
OY 422 DFG 424
DB 473 DFG 475

```

RESULT 3

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - rabbit
 N:Alternate names: tryptophanyl-tRNA synthetase
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 30-Sep-1992 #sequence, revision 13-Feb-1998 #text, change 26-May-2000
 C:Accession: A35904; S37396
 R:Lee, C.C.; Craig, W.T.; Muzny, D.M.; Harlow, E.; Caskey, C.T.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequ
 A:Reference number: A35904; MUID:90239043
 A:Accession: A35904
 A:Molecule type: mRNA
 A:Residues: 1-475 <LEE>
 A:Cross-references: GB:M33460
 R:Filova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegon, G.; McCaughan, K
 EMBO J. 12, 4013-4019, 1993
 A:Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase
 A:Reference number: S37396; MUID:94009008
 A:Accession: S37396
 A:Molecule type: mRNA
 A:Residues: 166-177 <FRO>
 C:Genetics:
 A:Gene: WRS
 C:Complex: homodimer [validated, MUID:94009008]
 C:Function:
 A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent for
 A:Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain rele
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

